

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
21 mai 2004 (21.05.2004)

PCT

(10) Numéro de publication internationale
WO 2004/041841 A2

(51) Classification internationale des brevets⁷ : **C07K**

(21) Numéro de la demande internationale :
PCT/FR2003/003293

(22) Date de dépôt international :
4 novembre 2003 (04.11.2003)

(25) Langue de dépôt : **français**

(26) Langue de publication : **français**

(30) Données relatives à la priorité :
02/13792 5 novembre 2002 (05.11.2002) **FR**

(71) Déposants (*pour tous les États désignés sauf US*) :
UNIVERSITE DE LA MEDITERRANEE (AIX-MARSEILLE II) [FR/FR]; Jardin du Pharo, 58, boulevard Charles Livon, F-13284 Marseille Cedex 07 (FR).
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) [FR/FR]; 3, rue Michel-Ange, F-75794 Paris Cedex 16 (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (*pour US seulement*) : **RAOULT, Didier [FR/FR]**; 16, rue de Lorraine, F-13008 Marseille (FR). **DRANCOURT, Michel [FR/FR]**; 9, Traverse de la Pauline, F-13012 Marseille (FR).

(74) Mandataire : **DOMANGE, Maxime**; Cabinet Beau de Loménie, 232, avenue du Prado, F-13295 Marseille Cedex 08 (FR).

(81) États désignés (*national*) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) États désignés (*regional*) : brevet ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Déclaration en vertu de la règle 4.17 :

— *relative à la qualité d'inventeur (règle 4.17.iv)) pour US seulement*

Publiée :

— *sans rapport de recherche internationale, sera republiée dès réception de ce rapport*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(54) Title: MOLECULAR IDENTIFICATION OF BACTERIA OF GENUS *STREPTOCOCCUS* AND RELATED GENUSES

(54) Titre : IDENTIFICATION MOLECULAIRE DES BACTERIES DU GENRE *STREPTOCOCCUS* ET GENRES APPARENTES

(57) Abstract: The invention concerns a method for detecting by molecular identification a bacterium of one of the species of genera *Streptococcus* and four related genera *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella* which consists in using as probe or primer: the *rpoB* gene or fragment of one said bacterium of sequences SEQ ID N° 1 to 3, or an oligonucleotide or a mixture of oligonucleotides derived from sequences SEQ ID N° 8 to 35, or in particular oligonucleotides of sequences SEQ ID n° 6 and 7.

(57) Abrégé : La présente invention a pour objet un procédé de détection par identification moléculaire d'une bactérie de l'une des espèces des genres *Streptococcus* et 4 genres apparentés *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella* pour lequel on utilise comme sonde ou amorce: -le gène ou fragment de gène *rpoB* d'une dite bactérie des séquences SEQ ID N° 1 à 3, ou - un oligonucléotide ou mélange d'oligonucléotides tiré des séquences SEQ ID N° 8 à 35, ou notamment les oligonucléotides des séquences SEQ ID n° 6 et 7.

WO 2004/041841 A2

BEST AVAILABLE COPY

IDENTIFICATION MOLECULAIRE DES BACTERIES DU GENRE *STREPTOCOCCUS* ET GENRES APPARENTES

La présente invention concerne le domaine du diagnostic. Plus précisément, l'invention concerne une méthode pour l'identification moléculaire des bactéries du genre *Streptococcus* et genres apparentés *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella* par les techniques de détection et/ou d'amplification et séquençage à l'aide de sondes ou d'amorces oligonucléotidiques appliquées à des souches de ces genres bactériens.

Les bactéries du genre *Streptococcus* et de quatre genres apparentés : *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella*, sont des bactéries cocciformes, gram positif et catalase négative dont on reconnaît actuellement plus d'une quarantaine d'espèces. Les bactéries du genre *Lactococcus*, précédemment classées parmi les streptocoques comme *Streptococcus* groupe N, n'entrent pas dans le champ de ce brevet du fait de leur rareté en pathologie humaine, et du fait qu'elles sont facilement discriminées des streptocoques par leur croissance à + 10°C. Le genre *Streptococcus* comporte officiellement 55 espèces. Le genre *Gemella* comporte 6 espèces, le genre *Abiotrophia* comporte 1 espèce, le genre *Granulicatella* comporte 3 espèces, le genre *Enterococcus* comporte 24 espèces [www.springer-ny.com/bergeysoutline/main.htm]. Ces espèces sont facilement et fréquemment cultivées à partir de prélèvements environnementaux, de prélèvements cliniques vétérinaires et de prélèvements cliniques humains [Ruoff KI. (1999) in *Manuel of Clinical Microbiology*, pp. 283-296, ASM press]. Chez l'homme, différentes espèces du genre *Streptococcus* sont responsables d'infections communautaires éventuellement sévères du fait du caractère invasif des streptocoques considérés ou du fait de la production de toxines et de manifestations cliniques éventuellement graves à distance du foyer infectieux. Par exemple, *Streptococcus pyogenes* (Streptocoque groupe A) est responsable d'angines et de syndromes post-streptococciques incluant le rhumatisme articulaire aigu au cours duquel la destruction des valves cardiaques par un processus inflammatoire est responsable d'une valvulopathie éventuellement mortelle. Également, plusieurs espèces du genre *Streptococcus* en particulier les Streptocoques du groupe A, du groupe C, et du groupe G sont

responsables d'infections invasives mortelles en particulier de myosite c'est-à-dire de destruction des tissus cutanés et sous cutanés et du tissu musculaire comme cela a été décrit depuis quelques années. Egalement par exemple *Streptococcus pneumoniae* (pneumocoque) est responsable de pneumonie, de 5 méningite et de septicémie. Par ailleurs, les bactéries des genres *Streptococcus*, *Enterococcus*, *Gemella*, *Abiotrophia*, et *Granulicatella* sont responsables d'endocardites c'est-à-dire d'infection des valves cardiaques chez l'homme, lesquelles constituent des maladies infectieuses mortelles [Casalta JP et al. Journal Clinical Microbiology, 2002, 40 : 1845-1847]. Egalement, certaines 10 espèces des genres considérés sont responsables d'infections nosocomiales, par exemple, les bactéries du genre *Streptococcus* du groupe A sont responsables de bactériémies qui succèdent à des explorations par endoscopie digestive. Egalement, les bactéries du genre *Enterococcus* sont responsables d'infections urinaires nosocomiales après utilisation d'antibio-prophylaxie par des 15 antibiotiques de la famille des céphalosporines auxquelles elles sont naturellement résistantes. Ces espèces bactériennes posent par ailleurs le problème de leur résistance croissante aux antibiotiques, résistance à la pénicilline G de *Streptococcus pneumoniae* [Garav J. Lancet Infect. Dis. 2002, 2 : 404-415] et résistance à la vancomycine d'*Enterococcus* spp. [Gold H.S. 20 Clin. Infect. Dis. 2001, 33 : 210-219 ; Bonten M.J. et al. Lancet Infect. Dis. 2001, 1 : 314-325].

Ces différentes espèces bactériennes posent le problème de leur détection dans les prélèvements pathologiques chez l'homme et de leur identification lorsqu'elles ont été isolées à partir desdits prélèvements. Les 25 méthodes conventionnelles de détection reposent en effet sur la mise en évidence de bactéries cocciformes gram positif, à l'examen direct du produit pathologique. Il est cependant connu que cette détection microscopique des bactéries du genre *Streptococcus* et de genres apparentés dans les prélèvements cliniques a un seuil de sensibilité de 10^4 CFU/ml. Il est donc tout à 30 fait possible qu'un prélèvement pathologique chez l'homme ou chez l'animal contienne une des espèces considérées qui ne soit pas détectée à l'examen microscopique direct de ce prélèvement pathologique. Par ailleurs, bien que leur structure soit celle de bactéries Gram – positif, elles peuvent apparaître

faussement Gram-négatif après coloration de Gram du prélèvement pathologique et donner lieu à une identification erronée ou à une impasse d'identification. Ceci est particulièrement fréquent pour les bactéries du genre *Gemella*. Chez l'homme, c'est en particulier le cas lors de l'examen 5 anatomopathologique et bactériologique des valves cardiaques dans le cas d'une endocardite.

Lorsque qu'une bactérie d'une espèce des genres considérés est isolée au laboratoire, les méthodes conventionnelles d'identification phénotypique sont les plus couramment utilisées pour l'identification des bactéries des espèces du 10 genre *Streptococcus* et des genres apparentés et plusieurs trousseaux d'identification ainsi que des automates ont été développés pour aider à l'identification phénotypique des bactéries du genre *Streptococcus* et des genres apparentés. Sur ce plan, le degré d'identification en pratique courante est variable. En particulier, un des tests utilisés pour l'identification des 15 streptocoques et des bactéries des genres apparentés est l'observation d'une réaction hémolytique, c'est-à-dire la destruction par la bactérie des hématies contenues dans une gélose au sang. Cependant cette réaction d'hémolyse peut être inhibée par la présence d'oxygène ou par la présence de peroxyde lorsque les bactéries streptocoques sont cultivées en présence de concentration 20 importante de dioxyde de carbone. Il est par ailleurs reconnu qu'il existe un certain degré de subjectivité dans l'appréciation de l'hémolyse par les colonies de streptocoques et donc une variabilité d'inter opérateur qui nuit ensuite à la qualité de l'identification de ces bactéries. Pour les streptocoques alpha-hémolytiques, un deuxième test est celui de la sensibilité à l'optochine qui 25 permet de reconnaître *Streptococcus pneumoniae* qui est sensible à ce composé. Cependant, des souches de *Streptococcus pneumoniae* résistant à l'optochine ont été rapportées [Lund E. Acta Patho. Microbiol. Immunol. Scand. 1959, 47, 308-315]. Un dernier test phénotypique est le sérotypage, ce test peut être faussement positif en particulier pour les streptocoques de sérogroupe D du fait d'antigénicité croisée entre les streptocoques du groupe D, *Enterococcus* et 30 *Pediococcus*.

Plusieurs systèmes moléculaires ont été développés pour l'identification de certains sérogroupes ou de certaines espèces du genre *Streptococcus*, en

particulier les streptocoques du groupes A (*Streptococcus pyogenes*, *Streptococcus anginosus*, *Streptococcus constellatus*, *Streptococcus intermedius*) et du groupe B (*Streptococcus agalactiae*) [Daly J.A. et al. *J Clin Microbiol.* 1991, 29 : 80-82 ; Heelan J.S. et al. *Diagn. Microbiol. Infect. Dis.* 1996, 24 : 65-69] de même que pour *Streptococcus pneumoniae* [Denys G.A. et Carrey R.B. *J. Clin. Microbiol.* 1992, 30 : 2725-2727] par hybridation de sondes spécifiques ciblant le gène codant l'ARN ribosomal 16S. Egalement, différents systèmes basés sur l'amplification par PCR de gènes codant pour des toxines ou des facteurs de virulence ont été développés pour la discrimination de *Streptococcus pneumoniae* parmi les Streptocoques α -hémolytiques [Salo P. et al. *J. Infect. Dis.* 1995, 171 : 479-482 ; Morrisson K. et al. *J. Clin. Microbiol.* 2000, 38, 434-437 ; Kaijalainen T. et al. *J. Microbiol. Meth.* 2002, 51 : 111-118], ainsi que pour la détection de *Streptococcus agalactiae* [Mawn J.A. et al. *J. Clin. Pathol.* 1993, 46 : 633-636]. Ces différents systèmes cependant ne permettent l'identification que d'une ou quelques espèces du genre *Streptococcus*.

Un système d'identification de trois espèces de streptocoque a été développé, basé sur l'amplification de l'entretoise 16S-23S [Forsman P. et al. *Microbiology*, 1997, 143, 3491-3500], mais l'identification n'a été limitée dans ce travail qu'à certaines espèces d'intérêt animal : *Streptococcus agalactiae*, *Streptococcus dysgalactiae* et *Streptococcus uberis*. Par ailleurs, il est actuellement indispensable de disposer dans les laboratoires de 2 cibles moléculaires distinctes pour la détection et l'identification des streptocoques, ceci afin de pallier les risques de contamination moléculaire inhérents à l'utilisation d'une seule cible.

Enfin, aucun système de détection et d'identification des genres apparentés à *Streptococcus* n'a été développé et plus particulièrement pour les bactéries du genre *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella*.

Les inventeurs ont démontré selon la présente invention, que le gène *rpoB* constitue un marqueur génétique permettant la détection et l'identification spécifique de la bactérie de chaque espèce du genre *Streptococcus* et de 4 genres apparentés : *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella*.

Bien que ce gène ait été précédemment montré comme un outil d'identification bactérienne dans différents genres bactériens, aucune publication

ne fait mention de son utilisation pour l'identification des bactéries des genres *Streptococcus* et des quatre genres apparentés et il n'y avait donc aucune suggestion quant à l'intérêt de la séquence de ce gène pour l'identification des dites bactéries. Au contraire, quelques séquences partielles du gène *rpoB* chez 5 quelques espèces, disponibles dans GenBank montrait une faible hétérogénéité, faisant douter de l'intérêt de ce gène comme outil d'identification pour ces bactéries. Enfin, les inventeurs ont développé un outil d'identification de quatre genres bactériens simultanément, obligeant la mise au point d'amorces dégénérées qui ne pouvaient être déduites d'aucune des séquences *rpoB* 10 déterminées pour chaque espèce.

Plus particulièrement, la présente invention concerne des séquences d'acides nucléiques spécifiques du genre ou de chaque espèce du genre *Streptococcus* et des genres apparentés dont la séquence nucléotidique est tirée du gène *rpoB* des dites bactéries.

15 Selon Lazcano et al. [J. Mol. Evol. (1988) 27 :365-376], les ARN polymérases sont divisées en deux groupes selon leur origine, l'un constitué par les ARN polymérases virales ARN- ou ADN-dépendantes, et l'autre constitué par les ARN polymérases ADN-dépendantes d'origine eucaryote ou procaryotes (archaébactéries et eubactéries). Les ARN polymérases ADN-dépendantes 20 eubactériennes sont caractérisées par une constitution multimérique simple et conservée notée « core enzyme », représentée par $\alpha\beta\beta'$, ou « holoenzyme » représentée par $\alpha\beta\beta'\sigma$ [Yura and Ishihama, Ann. Rev. Genet. (1979) 13 :59-97].

De nombreux travaux ont mis en évidence le rôle fonctionnel, au sein du complexe enzymatique multimérique, de la sous-unité β de l'ARN polymérase eubactérienne. Les ARN polymérases archaébactérienne et eucaryote présentent, pour leur part, une structure plus complexe pouvant atteindre une dizaine, voire une trentaine de sous-unités [Pühlet et al. Proc. Natl. Acad. Sci. USA (1989) 86 :4569-4573].

30 Les gènes qui codent les différentes sous-unités $\alpha\beta\beta'\sigma$ de l'ARN polymérase ADN-dépendante chez les eubactéries, respectivement les gènes *rpoA*, *rpoB*, *rpoC* et *rpoD*, sont classés en différents groupes comprenant les gènes codant pour des protéïnes constitutives des sous-unités ribosomiques ou pour des enzymes impliqués dans la réplication et la réparation du génome

[Yura and Yshihma, Ann. Rev. Genet. (1979) 13 :59-97]. Certains auteurs ont montré que les séquences des gènes *rpoB* et *rpoC* pouvaient être utilisées afin de construire des arbres phylogénétiques [Rowland et al. Biochem. Soc. Trans. (1992) 21 :40S] permettant de séparer les différents embranchements et sous-embranchements parmi les règnes du vivant.

Avant d'exposer plus en détail l'invention, différents termes, utilisés dans la description et les revendications, sont définis ci-après :

- Par « acide nucléique extrait de bactéries » on entend soit l'acide nucléique total, soit l'ADN génomique, soit les ARN messagers, soit encore l'ADN obtenu à partir de la transcription inverse des ARN messagers.
- Un « fragment nucléotidique » ou un « oligonucléotide » sont deux termes synonymes désignant un enchaînement de motifs nucléotidiques caractérisé par une séquence informationnelle des acides nucléiques naturels (ou éventuellement modifiés) et susceptibles de s'hybrider, comme les acides nucléiques naturels, avec un fragment nucléotidique complémentaire ou sensiblement complémentaire, dans des conditions prédéterminées de stringence stricte. L'enchaînement peut contenir des motifs nucléotidiques de structure différente de celle des acides nucléiques naturels. Un fragment nucléotidique (ou oligonucléotide) peut contenir par exemple jusqu'à 100 motifs nucléotidiques. Il contient généralement au moins 8, et en particulier au moins 12 motifs nucléotidiques, en particulier de 18 à 35, et peut être obtenu à partir d'une molécule d'acide nucléique naturelle et/ou par recombinaison génétique et/ou par synthèse chimique.
- Un motif nucléotidique est dérivé d'un monomère qui peut être un nucléotide naturel d'acide nucléique dont les éléments constitutifs sont un sucre, un groupement phosphate et une base azotée choisie parmi l'adénine (A), la guanine (G), l'uracile (U), la cytosine (C), la thymine (T) ; ou bien le monomère est un nucléotide modifié dans l'un au moins des trois éléments constitutifs précédents ; à titre d'exemple, la modification peut intervenir soit au niveau des bases, avec des bases modifiées telles que l'inosine, qui peut s'hybrider avec toute base A, T, U, C ou G, la méthyl-5-désoxycytidine, la désoxyuridine, la diméthylamino-5-désoxyuridine ou toute autre base modifiée capable d'hybridation, soit au

niveau du sucre, par exemple le remplacement d'un moins un désoxyribose par un polyamide [Nielsen PE et al., *Science* (1991) 254 :1497-1500], soit encore au niveau du groupement phosphate, par exemple par remplacement par des esters choisis notamment parmi les diphosphates, les alkylphosphonates et les phosphorothioates.

5

Par « hybridation », on entend le processus au cours duquel, dans des conditions appropriées, deux fragments nucléotidiques ayant des séquences suffisamment complémentaires sont susceptibles de s'associer par des liaisons hydrogène stables et spécifiques, pour former un double

10 brin. Les conditions d'hybridation sont déterminées par la « stringence », c'est à dire la rigueur des conditions opératoires. L'hybridation est d'autant plus spécifique qu'elle est effectuée à plus forte stringence. La stringence est fonction notamment de la composition en bases d'un duplex sonde/cible, ainsi que par le degré de mésappariement entre deux acides

15 nucléiques. La stringence peut également être fonction des paramètres de la réaction d'hybridation, tels que la concentration et le type d'espèces ioniques présentes dans la solution d'hybridation; la nature et la concentration d'agents dénaturants et/ou la température d'hybridation. La stringence des conditions dans lesquelles une réaction d'hybridation doit

20 être réalisée dépend notamment des sondes utilisées. Toutes ces données sont bien connues et les conditions appropriées peuvent éventuellement être déterminées dans chaque cas par des expériences de routine. En général, selon la longueur des sondes utilisées, la température pour la réaction d'hybridation est comprise entre environ 20 et 65°C, en particulier entre 35 et 65°C dans une solution saline à une concentration d'environ

25 0,8 à 1 M.

30 - Une « sonde » est un fragment nucléotidique possédant une spécificité d'hybridation dans des conditions déterminées pour former un complexe d'hybridation avec un acide nucléique ayant, dans le cas présent, une séquence nucléotidique comprise soit dans un ARN messager, soit dans un ADN obtenu par transcription inverse dudit ARN messager, produit de transcription ; une sonde peut être utilisée à des fins de diagnostic (notamment sondes de capture ou de détection) ou à des fins de thérapie,

- Une « sonde de capture » est une sonde immobilisée ou immobilisable sur un support solide par tout moyen approprié, par exemple par covalence, par adsorption, ou par synthèse directe sur un solide. Des exemples de supports comprennent les plaques de microtitration et les puces à ADN.
- 5 - Une « sonde de détection » est une sonde marquée au moyen d'un agent marqueur choisi par exemple parmi les isotopes radioactifs, les enzymes, en particulier les enzymes susceptibles d'agir sur un substrat chromogène, fluorogène ou luminescent (notamment une peroxydase ou une phosphatase alcaline), les composés chimiques chromophores, les
- 10 composés chromogènes, fluorogènes ou luminescents, les analogues des bases nucléotidiques et les ligands tels que la biotine.
- Une « sonde d'espèce » est une sonde permettant l'identification spécifique de l'espèce d'une bactérie.
- 15 - Une « sonde de genre » est une sonde permettant l'identification spécifique du genre d'une bactérie.
- Une « amorce » est une sonde comprenant par exemple 10 à 100 motifs nucléotidiques et possédant une spécificité d'hybridation dans des conditions déterminées pour les réactions d'amplification enzymatique.
- Par « réaction d'amplification » on entend une réaction de polymérisation
- 20 enzymatique, par exemple dans une technique d'amplification telle que la PCR, initiée par des oligonucléotides amores et utilisant une ADN polymérase.
- Par « réaction de séquençage », on entend l'obtention de la séquence d'un fragment d'acide nucléique ou d'un gène complet par un procédé de polymérisation abortive à partir d'amorces oligonucléotidiques et utilisant lesdits didésoxynucléotides (Sanger F, Coulson AR (1975), J.Mol.Biol. 94 : 441) ou hybridations multiples avec des sondes multiples fixées sur support solide telles qu'utilisées dans les puces ADN par exemple.
- 25 Les séquences des gènes *rpoB* des bactéries *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus mutans* et *Streptococcus agalactiae* ont été décrites dans la littérature.

Les inventeurs ont déterminé les séquences complètes des gènes *rpoB* d'autres espèces de bactéries du genre *Streptococcus* et apparentées:

Streptococcus anginosus et *Streptococcus equinus*, d'*Abiotrophia defectiva*, et une très large portion du gène pour *Streptococcus mutans* et *Enterococcus faecalis*. Ces espèces ont été choisies par les inventeurs comme représentant les principaux groupes génétiques déterminés sur la base de l'étude du gène 5 16S dans les bactéries du genre *Streptococcus* et genres apparentés, encadrant l'ensemble des espèces actuellement décrites dans ce genre, de sorte que l'alignement des séquences *rpoB* obtenues chez ces espèces puisse encadrer vraisemblablement l'ensemble des séquences *rpoB* de toutes les espèces de ces genres bactériens plus précisément, il s'agit donc des espèces 10 phylogénétiquement les plus divergentes parmi l'ensemble des espèces actuellement décrites dans ce genre, de sorte que l'alignement des séquences *rpoB* obtenu chez ces espèces puisse encadrer phylogénétiquement vraisemblablement l'ensemble des séquences *rpoB* de toutes les espèces de ce genre bactérien.

15 A partir de ces séquences complètes ou quasi complètes et après de nombreuses tentatives infructueuses tel que rapporté dans les exemples 1 et 2 ci-après, les inventeurs ont mis en évidence les séquences consensus et spécifiques SEQ.ID. n° 6 et 7 suivantes :

- SEQ ID N° 6 : 5' AARYTNGGMCCCTGAAGAAAT-3', et
- 20 - SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3',

dans lesquelles :

- N représente l'inosine ou l'un des 4 nucléotides A, T, C ou G,
- R représente A ou G,
- M représente A ou C, et
- 25 - Y représente C ou T,

et les séquences inverses et séquences complémentaires.

Les inventeurs ont déterminé lesdites séquences SEQ.ID.n°6 et 7 comme étant non seulement consensuelles entre toutes les bactéries du genre *Streptococcus* et desdits 4 genres apparentés mais en outre spécifiques de la 30 famille des bactéries du genre *Streptococcus* et desdits 4 genres apparentés.

A la position correspondant à un nucléotide N, Y, M ou R dans les séquences SEQ.ID. n°6 et 7 on trouve des nucléotides variables dans les séquences cibles complémentaires en fonction de l'espèce de la bactérie

considérée, mais tous les autres nucléotides sont conservés dans toutes les espèces des bactéries du genre *Streptococcus* et desdits 4 genres apparentés.

Des séquences SEQ.ID n°6 et 7 ainsi définies sont présentes dans les gènes *rpoB* de toute bactérie du genre *Streptococcus* et desdits 4 genres apparentés et spécifiques des bactéries du genre *Streptococcus* et desdits 4 genres apparentés et peuvent donc fournir des sondes de genre ou des amorces d'amplification pour détecter toute bactérie du genre *Streptococcus* et desdits 4 genres apparentés.

A cet effet, la présente invention a donc pour objet un oligonucléotide qui 10 comprend une séquence d'au moins 8, de préférence au moins 12, de préférence encore de 18 à 35, motifs nucléotidiques, dont au moins une séquence de 8, de préférence 12, de préférence encore 18 motifs consécutifs inclus dans l'une des séquences SEQ.ID. n° 6 et 7 suivantes :

15 - SEQ ID N° 6 : 5' AARYTNGGM CCTGAAGAAAT-3', et
- SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3',

dans lesquelles :

20 - N représente l'inosine ou l'un des 4 nucléotides A, T, C ou G,
- R représente A ou G,
- M représente A ou C, et
- Y représente C ou T,

et les séquences inverses et séquences complémentaires.

La présente invention a également pour objet un mélange d'oligonucléotides, caractérisé en ce qu'il est constitué d'un mélange équimolaire de d'oligonucléotides selon l'invention, ayant tous une séquence différente et 25 comprenant tous une séquence incluse dans SEQ ID n°6 ou tous une séquence incluse dans SEQ ID n°7.

Plus particulièrement, la présente invention a également pour objet un mélange de dits oligonucléotides, constitué d'un mélange équimolaire de 32 oligonucléotides de séquences différentes comprenant chacune au moins 15, de préférence encore au moins 18 motifs nucléotidiques consécutifs inclus dans la 30 séquence suivante :

- SEQ ID N° 6 : 5' AARYTNGGM CCTGAAGAAAT-3',
dans laquelle :

- R représente A ou G,
- Y représente C ou T,
- M représente A ou C, et
- N représente A, T, C ou G,

5 et les séquences inverses et séquences complémentaires.

La présente invention a également pour objet un mélange de dits oligonucléotides, constitué d'un mélange équimolaire de 8 oligonucléotides de séquences différentes comprenant chacune au moins 15, de préférence encore au moins 18 motifs nucléotidiques consécutifs inclus dans la séquence suivante :

10 - SEQ ID N° 6 : 5' AARYTNGGMCCCTGAAGAAAT-3',

dans laquelle :

- R représente A ou G,
- Y représente C ou T,
- M représente A ou C, et
- N représente l'inosine,

15 et les séquences inverses et séquences complémentaires.

La présente invention a également pour objet un mélange de dits oligonucléotides, constitué d'un mélange équimolaire de 16 oligonucléotides de séquences différentes comprenant chacune au moins 15, de préférence encore au moins 21 motifs nucléotidiques consécutifs inclus dans la séquence suivante :

20 - SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3',

dans laquelle :

- R représente A ou G, et
- N représente A, T, C ou G,

25 et les séquences inverses et séquences complémentaires.

La présente invention a également pour objet un mélange de dits oligonucléotides, constitué d'un mélange équimolaire de 4 oligonucléotides de séquences différentes comprenant chacune au moins 15, de préférence encore au moins 21 motifs nucléotidiques consécutifs inclus dans la séquence suivante :

30 - SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3',

dans laquelle :

- R représente A ou G, et
- N représente l'inosine,

et les séquences inverses et séquences complémentaires.

Lesdits mélanges d'oligonucléotides peuvent s'hybrider avec une séquence complémentaire incluse dans le gène *rpoB* de toutes les bactéries du genre *Streptococcus* et desdits 4 genres apparentés et peuvent donc être utilisés à titre de sonde de genre ou amorces d'amplification pour la détection ou respectivement l'amplification d'un fragment de gène *rpoB* d'une dite bactérie.

Pour préparer un dit mélange équimolaire d'oligonucléotides selon les synthèses d'oligonucléotides connues de l'homme de l'art, il suffit de mettre en œuvre un mélange équimolaire de 4 ou 2 nucléotides pour les nucléotides correspondant à N ou respectivement K, N, R ou Y, à savoir :

- un mélange équimolaire des 4 nucléotides, A, T, C et G pour les nucléotides correspondant à N dans lequel N représente A, T, C ou G, et
- un mélange équimolaire des 2 nucléotides T et G pour les nucléotides correspondant à K,
- 15 - un mélange équimolaire des 2 nucléotides A et C pour les nucléotides correspondant à N,
- un mélange équimolaire des 2 nucléotides A et G pour les nucléotides correspondant à R, et
- 20 - un mélange équimolaire des 2 nucléotides C et T pour un nucléotide représenté par Y.

On obtient ainsi un mélange équimolaire de 32 ($2^3 \times 4$) et 16 ($2^2 \times 4$) nucléotides de séquences différentes pour respectivement les 2 séquences SEQ ID n°6 et 7.

Dans lesdits mélanges équimolaires d'oligonucléotides selon l'invention, du fait que « N » représente l'inosine qui peut s'hybrider avec toute base ou un mélange équimolaire des 4 bases A, T, C, G, les séquences SEQ.ID.n° 6 et 7 peuvent s'hybrider avec la séquence complémentaire incluse dans le gène *rpoB* de toutes les bactéries du genre *Streptococcus* et desdits 4 genres apparentés.

En outre, ces séquences consensus SEQ.ID. n° 6 et SED ID n° 7 encadrent des séquences hyper variables dont la séquence est spécifique pour chaque espèce de bactérie du genre *Streptococcus*. Ces séquences encadrées par SEQ.ID. n° 6 et 7 peuvent donc être utilisées à titre de sonde d'espèce des bactéries du genre *Streptococcus* et desdits 4 genres apparentés.

De plus, les séquences SEQ.ID. n°6 et 7 ont été déterminées comme encadrant un fragment du gène *rpoB* comprenant une zone dont la longueur variable est d'environ 720 pb et comme comprenant les plus courtes séquences spécifiques pour chaque espèce de la bactérie du genre *Streptococcus* et 5 desdits 4 genres apparentés.

Les inventeurs ont ainsi pu mettre en évidence des sondes d'espèce pour chacune des 28 espèces de bactérie du genre *Streptococcus* et desdits 4 genres apparentés étudiées correspondant aux séquences SEQ.ID.n° 8 à 35 décrites à l'exemple 2 ci-après, encadrées par les séquences consensus SEQ.ID.n° 6 et 7.

10 Un autre objet de la présente invention est un gène ou fragment de gène *rpoB* d'une bactérie du genre *streptococcus* ou d'un desdits 4 genres apparentés, excepté les séquences SEQ ID n°11, 12, 14 et des bactéries *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* et *Streptococcus agalactiae*, les séquences inverses et séquences 15 complémentaires, caractérise en ce qu'il comprend une séquence telle que décrite dans les séquences SEQ ID n° 8 à 35 décrites à l'exemple 2;.

Un autre objet de la présente invention est la séquence complète du gène *rpoB* des bactéries, *Streptococcus anginosus*, *Streptococcus equinus* et *Abiotrophia defectiva* telles que décrites dans les séquences SEQ.ID. n° 1 à 3, 20 utiles notamment pour un procédé selon l'invention.

Un autre objet de la présente invention est la séquence quasi complète du gène *rpoB* de la bactérie *Enterococcus faecalis* telle que décrite dans la séquence SEQ.ID n° 5, utile notamment pour un procédé selon l'invention.

Dans les séquences SEQ.ID n° 1 à 3 et 5 et 8 à 35 décrites dans le 25 listage de séquences en fin de description :

- le nucléotide M représente A ou C,
- le nucléotide K représente T ou G,
- le nucléotide R représente A ou G,
- le nucléotide W représente A ou T,
- 30 - le nucléotide Y représente C ou T,
- le nucléotide N représente A, T, C, G ou I,
- le nucléotide S représente C ou G,
- le nucléotide V représente A, C ou G.

Les séquences consensus tirées de SEQ.ID.n° 6 et 7 mises en évidence selon la présente invention, peuvent être utilisées à titre de sonde de genre, d'amorce d'amplification ou de réaction de séquençage dans des procédés de détection de bactérie du genre *Streptococcus* et desdits 4 genres apparentés par 5 identification moléculaire.

..... Les séquences tirées des séquences SEQ.ID.n° 6 et 7 permettent donc non seulement de préparer des sondes de genre des bactéries du genre *Streptococcus* et desdits 4 genres apparentés mais aussi de détecter et identifier l'espèce de ladite bactérie par amplification et séquençage en utilisant lesdites 10 séquences comme amores.

La séquence complète du gène *rpoB* peut être utilisée pour identifier la bactérie pas seulement par l'étude de sa séquence primaire, mais aussi, par l'étude des structures secondaire et tertiaire de l'ARN messager provenant de la transcription de la séquence complète d'ADN.

15 Un autre objet de la présente invention est un oligonucléotide ou un fragment de gène *rpoB* ayant une séquence comprise ou consistant dans les séquences SEQ.ID.n°8 à 35, y compris donc les séquences SEQ ID n°11, 12, 14 et 22 des bactéries *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* et respectivement *Streptococcus agalactiae* et parmi les 20 oligonucléotides ou fragments de séquences inverses ou complémentaires tels que définis ci-dessus.

Les inventeurs, après analyse des différentes séquences, ont déterminé par comparaison deux à deux de toutes les séquences SEQ ID n°8 à 35, que, le taux d'homologie entre deux séquences différentes parmi lesdites séquences SEQ ID n°8 à 35 est au maximum de 98,7%. En dessous de 98,7% d'homologie 5 entre les séquences, celles-ci identifient des bactéries d'espèces différentes. En conséquence, la présente invention a également pour objet des oligonucléotides ou des fragments de gènes *rpoB* de séquences comprises ou consistant dans lesdites séquences SEQ ID n°8 à 35, les séquences inverses, les séquences complémentaires ainsi que dans les séquences présentant au moins 98,7% 10 d'homologie (c'est-à-dire un taux d'au moins 98,7% de similitude dans les séquences) par rapport aux dites séquences SEQ ID n°8 à 35, les séquences inverses et respectivement séquences complémentaires.

Les oligonucléotides, fragments de gène et gènes objets de la présente invention, ont été décrits comme comportant des séquences d'ADN, c'est-à-dire 15 avec des oligonucléotides A, T, C et G. Toutefois, la présente invention a également pour objet des oligonucléotides comprenant des séquences d'ARN correspondantes, c'est-à-dire dans lesquelles T est remplacé par U.

Dans la présente description, on entend par "séquences inverses et 20 séquences complémentaires", les séquences suivantes :

- la séquence inverse de ladite séquence,
- la séquence complémentaire de ladite séquence, et
- la séquence complémentaire de la séquence inverse de ladite séquence.

Les séquences SEQ.I. n°1 à 35 peuvent être préparées par génie 25 génétique et/ou par synthèse chimique, notamment par synthèse automatique, en utilisant les techniques bien connues de l'homme du métier.

Une première application d'un oligonucléotide selon l'invention est son utilisation comme sonde pour la détection, dans un échantillon biologique, de bactéries de l'une des espèces du genre *Streptococcus* et desdits 4 genres apparentés qui comprend une séquence nucléotidique dans l'une des 30 séquences SEQ.ID. n°6 à 35, et leurs séquences inverses ou complémentaires.

Un oligonucléotide comprenant les séquences SEQ.ID.n° 6 et 7 sera utilisé à titre de sonde de genre et un oligonucléotide comprenant une

séquence comprise dans ou comprenant l'une des séquences SEQ.ID. n° 8 à 35, sera utilisée à titre de sonde d'espèce.

Plus particulièrement, la présente invention a pour objet un oligonucléotide comprenant une séquence spécifique d'une espèce d'une bactérie du genre *streptococcus* et dits genres apparentés, de préférence d'au moins 20 nucléotides consécutifs, de préférence encore au moins 30 nucléotides consécutifs inclus dans l'une des dites séquences SEQ ID n°8 à 35, ou le cas échéant un mélange équimolaire de dits oligonucléotides de séquences différentes.

10 De préférence, lesdites séquences comprises dans l'une des séquences SEQ ID n°8 à 35, ayant de préférence au moins 20 nucléotides, de préférence encore au moins 30 nucléotides consécutifs inclus dans l'une des séquences SEQ ID n°8 à 35, constituent les séquences spécifiques des différentes espèces respectives les plus courtes, utilisables comme sonde d'espèces des bactéries 15 *Streptococcus* et des dits 4 genres apparentés concemées.

Les sondes selon l'invention peuvent être utilisées, à des fins de diagnostic, comme mentionné précédemment, par la détermination de la formation ou de l'absence de formation d'un complexe d'hybridation entre la sonde et un acide nucléique cible dans un échantillon, selon toutes les 20 techniques d'hybridation connues et notamment les techniques de dépôt ponctuel sur filtre, dites « DOT-BLOT » [Maniatis et al. (1982) Molecular Cloning, Cold Spring Harbor], les techniques de transfert d'ADN dites « SOUTHERN BLOT » [Southern E.M., J. Mol. Biol. (1975) 98 :503], les techniques de transfert d'ARN dites « NOTHERN BLOT », ou les techniques 25 dites « sandwich », en particulier avec une sonde de capture et/ou une sonde de détection, lesdites sondes étant capables de s'hybrider avec deux régions différentes de l'acide nucléique cible, et l'une au moins desdites sondes (généralement la sonde de détection étant capable de s'hybrider avec une 30 région de la cible qui est spécifique de l'espèce, étant entendu que la sonde de capture et la sonde de détection doivent avoir des séquences nucléotidiques au moins partiellement différentes.

L'acide nucléique à détecter (cible) peut être de l'ADN ou de l'ARN (le premier obtenu après amplification par PCR). Dans le cas de la détection d'une

cible de type acide nucléique double brin, il convient de procéder à la dénaturation de ce dernier avant la mise en œuvre du procédé de détection. L'acide nucléique cible peut être obtenu par extraction selon les méthodes connues des acides nucléiques d'un échantillon à examiner. La dénaturation 5 d'un acide nucléique double brin peut être effectuée par les méthodes connues de dénaturation chimique, physique ou enzymatique,...et en particulier par chauffage à une température appropriée, supérieure à 80°C.

Pour mettre en œuvre les technique d'hybridation précitées, et en particulier les techniques « sandwich », une sonde de l'invention, appelée sonde 10 de capture est immobilisée sur un support solide, et une autre sonde de l'invention, appelée sonde de détection, est marquée avec un agent marqueur. Les exemples de support et d'agent marqueur sont tels que définis précédemment.

De manière avantageuse, une sonde d'espèce est immobilisée sur un 15 support solide, et une autre sonde d'espèce est marquée par un agent marqueur.

Une autre application d'un oligonucléotide de l'invention est son utilisation comme amorce nucléotidique comprenant une oligonucléotide monocaténaire choisi parmi les oligonucléotides ayant une séquence d'au moins 12 motifs 20 nucléotidiques inclus dans l'une des séquences SEQ.ID. n° 6 à 35, qui est utilisable dans la synthèse d'un acide nucléique en présence d'une polymérase par un procédé connu en soi, notamment dans des méthodes d'amplification utilisant une telle synthèse en présence d'une polymérase (PCR, RT-PCR, etc.). En particulier, une amorce de l'invention peut être utilisée pour la transcription 25 inverse spécifique d'une séquence d'ARN messager de bactérie d'une espèce du genre *Streptococcus* et desdits 4 genres apparentés pour obtenir une séquence d'ADN complémentaire correspondante. Une telle transcription inverse peut constituer le premier stade de la technique RT-PCR, le stade suivant étant l'amplification par PCR de l'ADN complémentaire obtenu. On peut également 30 utiliser les amorces de l'invention pour l'amplification spécifique par réaction de polymérisation en chaîne de la séquence totale de l'ADN du gène *rpoB* d'une espèce du genre *Streptococcus* et desdits 4 genres apparentés.

Selon un cas particulier ladite amorce comprenant un oligonucléotide de l'invention comprend en outre la séquence sens ou anti-sens d'un promoteur reconnu par une ARN polymérase (promoteurs T7, T3, SP6 par exemple [Studier FW, BA Moffatt (1986) J. Mol. Biol. 189 :113] : de telles amorces sont 5 utilisables dans des procédés d'amplification d'acide nucléique faisant intervenir une étape de transcription, tels que, par exemple, les techniques NASBA ou 3SR [Van Gemen B. et al. Abstract MA 1091, 7th International Conference on AIDS (1991) Florence, Italy].

Un autre objet de l'invention est une amorce nucléotidique comprenant un 10 oligonucléotide choisi parmi les oligonucléotides ayant une séquence comprenant l'une des séquences SEQ ID n° 6 à 35 ou une séquence incluse dans SEQ.ID. n° 6 à 35 qui est utilisable pour le séquençage total ou partiel du gène *rpoB* d'une souche quelconque d'une espèce du genre *Streptococcus* et desdits 4 genres apparentés.

15 Le séquençage du gène *rpoB* partiel ou complet chez toute bactérie du genre *Streptococcus* et genres apparentés permet l'identification de toute bactérie *Streptococcus* et desdits 4 genres apparentés par analyse bio-informatique de cette séquence et la reconnaissance de nouvelles espèces de bactéries *Streptococcus* et desdits 4 genres apparentés inconnues.

20 De préférence, dans une utilisation comme amorce ou pour le séquençage des gènes *rpoB*, on utilise un dit mélange d'oligonucléotides selon l'invention, et de préférence encore des dits mélanges d'oligonucléotides consistant dans les séquences SEQ ID n°6 et SEQ ID n°7.

Plus précisément, la présente invention fournit un procédé de détection 25 par identification d'une bactérie de l'une des espèces du genre *Streptococcus* et desdits 4 genres apparentés caractérisé en ce qu'on utilise :

- un gène *rpoB* complet ou quasi complet de ladite bactérie selon la présente invention ainsi qu'un gène ou fragment de gène *rpoB* d'une bactérie *streptococcus pyogenes*, *streptococcus pneumoniae*, 30 *streptococcus mutans* et *streptococcus agalactiae* comprenant une séquence telle que décrite respectivement dans les séquences SEQ ID n°11, 12, 14 et 22, les séquences inverses et les séquences complémentaires, utiles notamment comme sonde d'espèces et/ou

- un dit fragment dudit gène *rpoB* de ladite bactérie selon la présente invention, comprenant une séquence nucléotidique choisie parmi l'une des séquences SEQ.ID.n° 8 à 35, les séquences inverses et les séquences complémentaires, utile notamment comme sonde d'espèces, et/ou
- un oligonucléotide selon la présente invention comprenant une séquence incluse dans l'une des séquences SEQ ID n°8 à 35, les séquences inverses et les séquences complémentaires, utile notamment comme sonde d'espèces, et/ou
- un oligonucléotide ou dit mélange d'oligonucléotides selon la présente invention comprenant une séquence constituée de motifs nucléotidiques consécutifs, inclus dans l'une des séquences SEQ.ID.n°6 et 7, utile notamment comme sonde de genre ou amorce d'amplification.

15 De préférence, dans ledit procédé de détection selon l'invention, on utilise :

- un dit fragment du gène *rpoB* de ladite bactérie comprenant une séquence choisie parmi l'une des séquences SEQ ID n° 8 à 35 ou un oligonucléotide de séquence comprise dans l'une des dites séquences SEQ ID n°8 à 35, les séquences inverses et séquences complémentaires, et/ou
 - au moins un dit mélange d'oligonucléotides selon la présente invention, dont les séquences de préférence consistent dans les séquences SEQ ID n° 6 et 7, et leurs séquences inverses et séquences complémentaires dans lesquelles de préférence encore N représente l'inosine.

20 25 Dans un premier mode de réalisation d'un procédé de détection selon l'invention, on cherche à mettre en évidence la présence d'une bactérie du genre *Streptococcus* et desdits 4 genres apparentés, on réalise les étapes dans lesquelles :

1. on met en contact au moins une sonde de genre comprenant un dit mélange d'oligonucléotides de séquences comprenant ou comprises dans l'une des séquences SEQ.ID.n° 6 et 7, les séquences inverses ou les séquences complémentaires selon l'invention, avec un échantillon contenant ou susceptible de contenir des acides nucléiques d'au moins

une telle bactérie du genre *Streptococcus* et desdits 4 genres apparentés, et

5 2. on détermine la formation ou l'absence de formation d'un complexe d'hybridation entre ladite sonde de genre et les acides nucléiques de l'échantillon, et on détermine la présence d'une dite bactérie du genre *Streptococcus* et desdits 4 genres apparentés s'il y a formation d'un complexe d'hybridation.

Dans une deuxième mode de réalisation d'un procédé de détection d'une bactérie du genre *Streptococcus* et desdits 4 genres apparentés, on réalise les 10 étapes dans lesquelles :

1. On met en contact des amores d'amplification comprenant desdits mélanges d'oligonucléotides comprenant une séquence incluse dans lesdites séquences SEQ.ID. n° 6 et 7, séquences inverses et séquences complémentaires selon l'invention, avec un échantillon 15 contenant ou susceptibles de contenir des acides nucléiques d'au moins une telle bactérie du genre *Streptococcus* et desdits 4 genres apparentés avec :

20 - comme amorce 5' : un dit mélange d'oligonucléotides comprenant une séquence incluse dans la séquence SEQ.ID.n° 6, ou de préférence consistant dans ladite séquence SEQ ID n° 6 complète, ou une séquence complémentaire selon l'invention ;

25 - comme amorce 3' : un dit mélange d'oligonucléotides comprenant une séquence incluse dans la séquence la séquence SEQ.ID.n° 7 ou de préférence consistant dans ladite séquence SEQ ID n° 7 complète, ou respectivement une séquence complémentaire selon l'invention.

30 2. On réalise une amplification d'acides nucléiques par réaction de polymérisation enzymatique et on détermine l'apparition ou l'absence d'un produit d'amplification, et on détermine ainsi la présence d'une dite bactérie dans l'échantillon si un produit d'amplification est apparu.

Ce deuxième mode de réalisation peut être utilisé pour détecter spécifiquement le genre d'une bactérie du genre *streptococcus* ou desdits 4 genres apparentés.

Mais, à l'étape 2 de ce deuxième mode de réalisation, on peut chercher à détecter spécifiquement une espèce donnée d'une bactérie du genre *Streptococcus* choisie parmi les espèces *Streptococcus mutans*, *Streptococcus oralis*, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus 5 salivarius*, *Streptococcus sanguinis*, *Streptococcus suis*, *Streptococcus acidominimus*, *Streptococcus agalactiae*, *Streptococcus anginosus*, *Streptococcus constellatus*, *Streptococcus difficile*, *Streptococcus dysgalactiae*, *Streptococcus equi*, *Streptococcus equinus*, *Streptococcus intermedius*, *Streptococcus mitis*, *Streptococcus bovis*, *Granulicatella adjacens*, *Abiotrophia 10 defectiva*, *Enterococcus avium*, *Enterococcus casselliflavus*, *Enterococcus faecalis*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus sacharolyticus*, *Gemella haemolysans* et *Gemella morbillorum*., comme décrit dans la variante de réalisation d'un procédé de détection spécifique d'une espèce de dites bactéries, décrite ci-après

15 Comme cela a été précédemment exposé en introduction, les genres *Streptococcus*, *Enterococcus*, *Granulicatella*, *Abiotrophia*, et *Gemella* comportent plus d'espèces bactériennes que celles qui ont été effectivement séquencées dans ce travail. Toutefois, les espèces séquencées ont été choisies telles qu'elles encadrent toutes les espèces connues dans ces genres bactériens 20 et sont en nombre suffisant pour démontrer l'application de la séquence *rpoB* à l'identification des espèces de ces genres.

Dans une variante de réalisation d'un procédé de détection spécifique d'une espèces desdites bactéries selon l'invention, on réalise les étapes dans lesquelles :

25 1- on met en contact un échantillon contenant ou susceptible de contenir des acides nucléiques d'au moins une telle bactérie, avec au moins une sonde d'espèce consistant dans un dit gène, dit fragment de gène ou dit oligonucléotide comprenant une séquence incluse dans l'une des séquences SEQ.ID n° 8 à 35, de préférence un oligonucléotide consistant dans l'une desdites séquences 30 SEQ.ID. n° 8 à 35, les séquences inverses et séquences complémentaires selon l'invention, et

2- on détermine la formation ou l'absence d'un complexe d'hybridation entre ladite sonde et les acides nucléiques de l'échantillon, et on détermine ainsi

la présence de ladite bactérie dans l'échantillon s'il y a formation d'un complexe d'hybridation.

Dans une autre variante de réalisation du procédé selon l'invention dans lequel on cherche à détecter spécifiquement une espèce donnée d'une bactérie 5 du genre *Streptococcus* et desdits 4 genres apparentés choisis parmi les 28 espèces citées ci-dessus, le procédé comprend les étapes dans lesquelles, dans un échantillon contenant ou susceptible de contenir des acides nucléiques d'au moins une dite bactérie :

- a) on réalise une réaction de séquençage d'un fragment du gène *rpoB* 10 amplifié d'une dite bactérie donnée à l'aide des amorces nucléotidiques consistant dans desdits mélanges d'oligonucléotides comprenant des séquences incluses dans la séquence SEQ.ID. n° 6 comme amorce 5', et SEQ.ID. n° 7 comme amorce 3', de préférence les séquences consistant dans lesdites séquences SEQ.ID. n° 6 et 7, et leurs 15 séquences complémentaires, et
- b) on détermine la présence ou l'absence de l'espèce donnée de ladite bactérie en comparant la séquence dudit fragment obtenu avec la séquence du gène complet *rpoB* de ladite bactérie ou la séquence d'un fragment du gène *rpoB* de ladite bactérie comprenant lesdites 20 séquences n° 8 à 35 et séquences complémentaires selon l'invention, et on détermine ainsi la présence de ladite bactérie dans l'échantillon si la séquence du fragment obtenue est identique à la séquence connue du genre ou du fragment de gène *rpoB* de ladite bactérie.

La présente invention a également pour objet une trousse de 25 diagnostic utile dans un procédé selon l'invention comprenant au moins un dit fragment de gène ou dit oligonucléotide de séquence comprise ou consistant dans les séquences SEQ.ID. n° 8 à 35 ou un dit oligonucléotide ou mélange d'oligonucléotides comprenant une séquence incluse dans une des séquences SEQ.ID. n° 6 et 7, et/ou au moins un dit fragment de gène *rpoB* d'une dite 30 bactérie comprenant les séquences SEQ.ID. n° 8 à 35, et les séquences complémentaires selon l'invention.

Avantageusement, un trousse selon la présente invention comporte des dits oligonucléotides sous forme de "biopuces", c'est-à-dire fixés sur des

supports solides, notamment en verre, selon le procédé décrit dans le brevet US 5 744 305 (Affymetrix, Fodor et al) en utilisant la stratégie de reséquençage décrite dans la demande WO 95/11995 (Affymax, Chee et al) ou selon la méthode décrite par A Troesch et al. dans J. Clin. Microbiol., vol. 37(1), p 49-55, 5 1999. Les oligonucléotides synthétisés sur la "biopuce" réalisent le reséquençage de la région hyper variable du gène *rpoB*. Ce procédé présente un avantage considérable en terme de coût de production et sans compromis sur la qualité de l'identification des différentes espèces de part le choix de ces séquences d'identification. De préférence, ces oligonucléotides fixés sur le 10 support solide de la "biopuce" comportent de 10 à 30 bases par exemple 20 bases, avec une position d'interrogation située dans la région centrale comme par exemple en 12ème position par rapport à l'extrémité 3' de la séquence pour des oligonucléotides de 20 bases. Un autre exemple consiste à utiliser des oligonucléotides de 17 bases, avec 2 positions d'interrogations : une en 10ème 15 et une en 8ème position. D'autres oligonucléotides ont des longueurs comprises entre 10 et 25 nucléotides. Les positions d'interrogation varient alors en fonction de la longueur de l'oligonucléotide:

L'analyse est effectuée sur le système complet GeneChip® (référence 900228, Affymetrix, Santa Clara, CA) qui comprend le lecteur GeneArray®, le 20 four d'hybridation GeneChip®, la station fluidique GeneChip® et le logiciel d'analyse GeneChip®.

Un oligonucléotide selon l'invention peut aussi être utilisé à titre de sonde de thérapie génique pour traiter les infections provoquées par une souche appartenant à une espèce du genre *Streptococcus* et desdits 4 genres 25 apparentés, ladite sonde comprenant un oligonucléotide tel que défini précédemment. Cette sonde de thérapie génique, capable de s'hybrider sur l'ARN messager et/ou sur l'ADN génomique desdites bactéries, peut bloquer les phénomènes de traduction et/ou transcription et/ou de réplication.

Le principe des méthodes de thérapie génique est connu et repose 30 notamment sur l'utilisation d'une sonde correspondant à un brin anti-sens : la formation d'un hybride entre la sonde et le brin sens est capable de perturber au moins l'une des étapes du décryptage de l'information génétique. Les sondes de thérapie génique sont donc utilisables comme médicaments antibactériens,

permettant de lutter contre les infections causées par les bactéries des espèces du genre *Streptococcus* et desdits 4 genres apparentés.

L'invention sera mieux comprise à l'aide de l'exposé ci-après, divisé en exemples, qui concernent des expériences effectuées dans le but de réaliser 5 l'invention et qui sont données à titre purement illustratif.

La figure 1. représente la visualisation des produits d'amplification par coloration au bromure d'éthidium après électrophorèse sur un gel d'agarose obtenu à l'exemple 3.

10 Exemple 1. Séquence du gène *rpoB* de trois espèces du genre *Streptococcus* et genre apparenté : *Abiotrophia defectiva*, *Streptococcus anginosus* et *Streptococcus equinus*.

La séquence complète du gène *rpoB* des bactéries des espèces *Abiotrophia defectiva*, *Streptococcus anginosus* et *Streptococcus equinus* a été 15 déterminée par amplification enzymatique et séquençage automatique disponible chez les Streptocoques. Le choix de ces espèces a été basé sur l'analyse de l'arbre 16S qui montre une divergence génétique couvrant l'ensemble de l'arbre phylogénétique des streptocoques.

Stratégie et Séquençage :

20 Plusieurs séquences partielles de 510 pb de gènes *rpoB* sont disponibles sur GenBank pour les 10 espèces de streptocoques suivantes : *Streptococcus intermedius*, *Streptococcus sanguinis*, *Streptococcus pneumoniae*, *Streptococcus parasanguinis*, *Streptococcus oralis*, *Streptococcus mitis*, *Streptococcus cristalus*, *Streptococcus constellatus*, *Streptococcus anginosus* et 25 *Granulicatella adjacens*. [Majewski,J., Zawadzki,P., Pickerill,P., Cohan,F.M. and Dowson,C.G. Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation.J. Bacteriol. 182, 1016-1023 (2000)], mais les amores utilisées par ces auteurs n'amplifient qu'une fraction des espèces du genre *Streptococcus* et il n'a donc pas été possible de mener à bien 30 notre travail sur la base de ces seules données. Il a donc fallu déterminer des amores capables d'amplifier l'ensemble des souches de streptocoques, enterocoques, *Abiotrophia*, *Gemella*, et *Granulicatella*. Ces amores devaient en outre encadrer une région présentant une diversité génétique suffisante pour

permettre de distinguer deux espèces entre elles. Cependant, l'alignement de ces séquences partielles publiées a permis de déterminer les amorces communes suivantes : (la numération se réfère à la séquence complète du *Streptococcus. pyogenes*)

5 SEQ ID N° 36 : 5'- AGACGGACCTTCTATGGAAAA -3' (amorce 748F) ,
- SEQ.ID.N°.37 : 5'- GGACACATACGACCATAGTG -3' (amorce 116R), et
SIQ N° 38 : 5'- GTTGTAACCTTCCAWGTCAT -3' (amorce 830R).

Ces amorces ont permis de séquencer la partie centrale du gène *rpoB* de 714pb pour les cinq espèces choisies (*Streptococcus equinus*, *Streptococcus mutans*, *Streptococcus anginosus*, *Enterococcus faecalis*, et *Abiotrophia defectiva* A partir de ce fragment central, le séquençage a été poursuivi par la technique dite du génome Walker.

En dehors de cette zone publiée [Majewski,J., et al. J. Bacteriol. 2002, 182, 1016-1023], l'alignement des deux séquences complètes disponibles dans 15 GenBank (*Streptococcus pneumoniae* [GenBank numéro d'accès AE008542] et *Streptococcus pyogenes* [GenBank numéro d'accès AE006480) ont permis de choisir les amorces suivantes :

- SEQ ID N° 39 : 5'- GTCTTCWTGGGYGATTTCCC -3' (amorce 2215R),
- SEQ ID N° 40 : 5'- ACCGTGGIGCWTGGTTRGAAT -3' (amorce 2057R),
20 - SEQ ID N° 41 : 5'- AACCAATTCCGYATYGGTYT -3'(amorce 1252R),
- SEQ ID N° 42 : 5'- AGIGGGTTAACATGATGTC -3'(amorce 371F),
- SEQ ID N° 43 : 5'- AGIGCCCAACCTCCATCTC -3'(amorce 730F), et
- SEQ ID N° 44 : 5'- CTCCAAGTGAACAGATGTGTA -3'(amorce 585R).

Ces amorces ont permis d'étendre la région séquencée pour certaines des cinq 25 souches choisies. De façon tout à fait inattendue, *E. faecalis* n'est pas amplifiée par ces amorces ; mais on a observé que la zone partielle séquencée présentait une homologie avec le gène *rpoB* de *Listeria monocytogenes*, c'est à dire avec une bactérie appartenant à un genre bactérien différent ce qui ne pouvait absolument pas être déduit des données existantes, on a donc choisi des 30 amorces dans le gène *rpoB* de *Listeria* pour amplifier le gène *rpoB* de *Enterococcus faecalis*.

- SEQ ID N°45 : 5'- TTACCAAACTTAATTGAGATTCAAAC- 3' (amorce 180F)
- SEQ ID N°46 : 5'- AGTATTTATGGGTGATTCCCA- 3' (amorce 410F)

- SEQ ID N°47 : 5'- GGACGTTATAAAAATCAACAAAAAATT- 3' (amorce 910F)
- SEQ ID N°48 : 5'- AGTTATAACCATCCCAAGTCATG- 3' (amorce 2430R)
- SEQ ID N°49 : 5'- TGAAGTTTATCATCAACCATGTG- 3' (amorce 3280R)
- SEQ ID N°50 : 5'- CCCAAAACGTTGCCACC- 3' (amorce 3360R)

5 Les séquences partielles ainsi obtenues pour les cinq souches choisies (*Streptococcus equinus*, *Streptococcus mutans*, *Streptococcus anginosus*, *Enterococcus faecalis*, *Abiotrophia defectiva*) ont permis de choisir les amores suivantes :

- SEQ ID N°51 : 5'- AACCAAGCYCGGTTAGGRAT -3' (amorce 520R)
- 10 - SEQ ID N°52 : 5'- ATGTTGAACCCACTIGGGGTGCCAT -3' (amorce 2881F) pour le séquençage des zones C- et N- terminales par Génome Walker.

15 Le séquençage est alors complet, comme en témoignent la détermination de la région codante, et l'alignement des protéines traduites des séquences nucléotidiques avec les deux protéines RpoB publiées de *Streptococcus pneumoniae* et *Streptococcus pyogenes*.

Plusieurs amores consensus potentielles ont fait l'objet d'investigations pour obtenir un fragment susceptible de conduire à la séquence complète des gènes *rpoB* par elongations successives à partir d'une série d'amores spécifiques.

20 Dans chacune des étapes ci-dessus, un grand nombre de tentatives avec des amores théoriquement ou potentiellement appropriées ont échoué avant de déterminer les amores mentionnées ci-dessus pour permettre d'amplifier et de séquencer par étapes successives la totalité des gènes *rpoB* décrite ci-après.

25 Les réactions de séquençage ont été réalisées en utilisant les réactifs du kit ABI Prism dRhodamine Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer Applied Biosystems) selon les recommandations du fournisseur suivant le programme suivant : 30 cycles comprenant une étape de dénaturation à 94°C pendant 10 sec., une étape d'hybridation de l'amorce à 50°C pendant 10 sec. et une étape d'extension à 60°C pendant 2 minutes. Les produits de 30 séquençage ont été séparés par électrophorèse sur un gel de polyacrylamide sur un séquenceur 377 DNA Sequencer (Perkin) et analysés pour former des séquences consensus par le logiciel Sequence Assembler (Applied Biosystems).

Cette approche nous a permis de déterminer la séquence complète du gène *rpoB* chez deux espèces du genre *Streptococcus* et chez *Abiotrophia defectiva* :

SEQ.ID. n°1 : Séquence du gène *rpoB* de *Streptococcus anginosus*. Cette séquence mesure 4 523 paires de bases , possède un contenu en cytosine plus guanosine de 41 % et est déposée dans GenBank sous le numéro d'accesion AF 535183 :

5'-TCATACTTTAGAGTCAGATTAGCTGCTCTTTGTGCCTGTTGGGATTTTGTGCGTTGT
 10 CATCAAAATTAAAGATTCTGAAAATTACTCAAAAAGGATAATGAAAATTGCTACTCTATTCCA
 TTAATAGAGAAATGTAGAAAAGAAGAAGGAGTAAAAAAACTTGGCAGGACATGAAGTTCAATACGGG
 AAACACCGTACTCGTCGTAGTTTCAAGAATCAAGGAAGTTCTGATTACCAAATTGATTG
 AAATCCAGAGGATTCTGTTCAAAGATTTCTGACCATGGTTGAAAGAAGTATTGAAGATGTA
 CTTCCCTATCTCAAACCTTACAGATAACAATGGAGCTAGAGTTGTGGTTATGAAATTAAAGGAT
 CTAAATACACTTTAGAAGAACGTATCCATGATGCCAGCTATTCTGCACCTATTGTGAC
 15 TTTCCGTTGATTAATAAGAAACTGGTCAAATCAAAACCAAGAAGTGTCTTGGCGATTTC
 CCAATCATGACAGAAATGGGAACTTCAATTATCAATGGTGGTGAGCGGATTATCGTATCTCAGC
 TCGTTCGTTCTCCAGGTGTTACTTCACGATAAGTAAAGTAGACAAAAATGGTAAAGTGGTTATGG
 TTCAACTGTCATTCTAACCGTGGAGCTGGTAGAGCTGGAAACAGACTCAAAAGATATTGCT
 TATACTCGGATTGACCGTACTCGTAAGATTCCGTTACGACACTTGTGCGCTGGTTTT
 20 CTGGCGATGATGAAATCTTGACATTTCGGCGACAGCGATCTCGTTGCCAACACGATTGAAAA
 GGATATTCAATAAAATCCAATGGATTCACTGACGGATGAAGCGCTAAAGAAATCTATGAACGT
 CTTCGTCCAGGTGAGCCTAAACAGCTGATAGTTCACTGAGTCTATTGGTCGCTCGTTCTTG
 ATCCACATCGTTACGACTTGGCGGAGTTGGTCTGGTATAAAATCAATAAAATTAACATTAA
 AACACGTTGTTAAATCAAACGATTGACAGGCCTTGGTAGATCCAGAAACAGGTGAAATCTG
 25 GTTGAAGCTGGAACGGTTATGACCGTAGTGTCAATTGATAGCATTGCAAGAAACTTGGACGGTG
 ATTTGAATAAAATCACTTATATTCCAATGATGACGCTGTGTTAACAGAGGCCAGTGTCTTCA
 AAAATTCAAAGTGGTGGCGCCAATGATCCAGATCGTGTGGTAGCTATTATTGGTAATGCCAAC
 CCAGGAGATCGAGTTACGATTACGCCAGCAGATATTGGCTGAGATGAATTACTTCTTGA
 ACCTCGCTGAAGGACTTGGTCGTGGACGATATTGACCACTTGGAAATCGTCGGATTCTG
 30 CGTTGGTGAATTGCTTGCTAACCAAGTACGTCTGGCTTGTCTCGTATGGAGCGAACGTTG
 GAGCGCATGAGTGTGCAAGATAATGAAGTGTGACACCAGCAACAAATCATTAACATCCGCCAG
 TCACAGCAGCTATCAAAGAATTCTTGGTCATCTCAATTGTCTCAATTATGGACCAACATAA
 TCCACTGTCGATTGTCTCACAAACGCCGTTGTCAGCCTGGGACCTGGTGGTTGACTCGT
 GATCGTGTGGATATGAAGTGCCTGACGTGCACTATACCCACTATGGCTGATGTGTCCGATTG
 35 AAACGCCCTGAAGGACCAACATCGTTGATCAATAACTGTCTTATGGACACTTGAATAA
 ATATGGCTTATCCAACGCCGTATCGTAAAGTGGATCGGAAACAGGTCTGGTCACCAATGAA
 ATCGTTGGCTGACAGCGGACGAAGAAGATGAATTATCGTAGCGCAAGCAAATTCTAAATTAA
 CAGAAGATGGTCGTTTGCAGAAGCGATTGTCAATGGACGTCACCAAGGGAAACAACCAAGAATT
 40 TCCTTCAGATCAAGTAGACTTCATGGATGTATCGCCTAACGAGGTAGTTGCGGTTGCGACAGCA
 TGTTATTCTTCTTGAAACGACGACTCAAACCGTGCTCTCATGGGTGCCAACATGCAACGTC
 AGGCGGTACCGTTGATTGATCCGCATGCCACCATATGGTACTGGTATGGAATACCAAGCAGC
 TCATGACTCTGGTGCAGCGATTATTGCCAACACGACGGTAAAGTTGTATATTGATGCAAGCC
 AAAGTTGAAGTTCGTCGTGAAGATGGCTACTTGATGTCTATCATATTACGAAATTCCGCCGTT
 CAAACTCTGGTACTTCTTACAACCAACGTACGCTGGTAAAAGTTGGCGATACAGTTGAAAAAGG
 45 TGACTTTATCGCAGACGGACCTCTATGGAAAAAGGTGAAATGGCACTTGGACAAAATCCAATC
 GTTGCTTATATGACATGGGAAGGTTACAACATTGAAAGATGCCGTTATCATGAGTGTGAGCGTTAG
 TGAAAGACGATGTTACACATCTGTCACTTGGAGGAATTGAATCAGAAACACGTGATACAAA STRF
GCTTGGACCTGAAGAAATCACGCCGAAATTCCAACACGTGGTGAAGATGCTTGTGAGAGACCTT
GACGAAACGGGAATTATCCGATTGGTGTGAGGTAAAAGAAGGCGACATTCTGTCGGTAAAG
 50 TAACACCGAAAGGTGAAAAGACTTATCTGCTGAAGAACGCCCTGCTTATGCAATTTCGGTGA
 TAAATCTCGTGAAGTACGTGATACTTCCCTTCGTGTACACATGGTGGTGCAGGGGTTGTCG
 GATGTGAAAATCTTACTCGTGCACGGTGTGAATTGCAATCTGGTGTCAACATGTTGGTAC
 GTGTTACATCGCTCAAAACGGAAAATCCGTGTGGGGATAAGATGGCTGGACGTCACGGAAA
 CAAAGGGTTGTTCCCGATTGTCAGTTGAGGATATGCCGTATCTCCAGATGGAACACCA

GTTGATATTATGTTGAACCCACTTGGGGGCCATCTGTATGAATATTGGTCAAGTTATGGAGC
 TTCACCTCGGTATGGCTGCTCGAACCTTGGCATTACATGCAACACCAAGTATTGACGGGGC
 TAGCTCAGATGATCTTGGGAAACCGTTGCTGAAGCTGGCATGGATAGCGATGCTAACAGAACATC
 CTTTATGATGCCGTACTGGTGAGCCATTGATAATCGTGTATCCGTGGTGTATGTACATGA
 5 TCAAACCTCCACCATATGGTGATGATAAGCTCCATGCCCGTTCCGTTGGCCTTATTCAACCGT STRR
 TACGCAACAACCTCTTGGTGGTAAAGCGCAGTTGGTGGACAACTGTTGGAGAAATGGAAGTT
 TGGGCTCTTGAAGCCTACGGTGTCTAACGTCTAACAGAAATCTGACTAACAGTCAGATG
 ACATCAATGGTCGTTGAGAGCTTATGAAGCCATTACCAAAGGTAAGCCAATTCCAAAACCAGG
 TGTTCCAGAACCTCCGTGTCTTGTAAAAGAATTGCAATCACCTGGTCTTGACATGCGTGTC
 10 CTTGATGAAGACGACAATGAAGTCGAACCTCGTACTTGGACGAAGGATGGATGATGATGTGA
 TTGATGTAGACGATCTGAAAAGCACGTGAAAAGCAGCACAAAGAAGCAAAAGCCGCTTTGA
 TGCTGAAGGAAAGAATAAGAACTGATTCAATAGATAATAAGAAAGGTAAGAAATAGTGGTTG
 ATGTAATCGTTCAAAGTATGCAAATCACCTAGCTTCTCTAGTAAAGTCCGCTTGGTC
 15 TTATGGAGAAGTGAAGAAACCTGAAACAATTAACTACCGCACACTAAAACCAGAACCGCAAGGG
 CTTTTGATGAAGTCATCTTGGCCTACGAAAGACTGGGAATGTCGTGGAAAATATAAAC
 GGATTCGTTATAAAGGAATCATTGTGACCGTTGGTGTGAAGTAACCTGACTAAAGTCG
 TCGTGAACGTATGGACATATTGAGTTGAAAGCCCCAGTCTCCTCATATTGGTATTTAAAGG
 AATTCCAANTCGCATGGCTTGACCTGGACATGAGCCCTCGTGTCTTGAAGAAGTCATNTAN
 20 TTTGCAGCTTATGTGGTGANTGACCTAAAGATAACNCCACTTGAGCACAAATCCATTATGACAG
 AGCGGGATGGTTNGTGAACGCTGACNTGAATATGCCAAGGCTTTGCAAAAATGGGTG
 YTGAAGCAATCCAAGATCCTGAAACANGTAGACCTGGAAAAGAAATTGCAGAGCTCAAAGA
 TGAATTAAAACGGCAAGTGGCAAAAGCGCGTAAAMGCTAANTCGTCGNTNNGACTCTTTC
 GATNCTTCCAAAATCATGGTACACAAAACCAGAACTGGATGGTCTTAAACCACNTNTCACC
 GCTCATTCCAGACAC -3'

25

SEQ ID N°2 : Séquence du gène *rpoB* de *Streptococcus equinus*. Cette séquence mesure 4 118 paires de bases et possède un contenu en cytosine plus guanosine de 41 % est déposée dans GenBank sous le numéro GenBank accession AF 535187:

30 5' -CACCGTGGTCGACGGCCCAGGGCTGGTGAATTGTCATAAGTTGTAGTAGTAAATTCCCTTAT
 CAGTGTGATGCATGAGCTATAAATAGTGTACTCATATTGCCACTTCATCGACATAGCAAAG
 TCCTTTTGTGTTCAACGGATTAAAATGTGGAAGAATTGATTAACACTGCTTCTTCTGTT
 TCTTCAGCCACAGAATTAAATTGTAAAAGTAACCTTACATAACGTGACATTGATGATAAT
 CACCAGGCAAGCCAAGTCCACCCATGCCACGGCTATAAGTTCAAGTTCTAACTCTTAGCAAA
 35 ACGATTTCCTGAAACCTTGGAGATAGATGACGATAGTTATTCAAATTGAATAATTGTTATCA
 AAAGTTGGATTATTAGTCAAAACACCTGTTGAGTTATTGCTAAACTTATAGGGCACCGTGGTC
 GACGGCCGGGCTGGTAAAGACTCTTGGATAACGGATTAAMAGAAGTTTGAAAGATGTACTT
 CCGATTACAAACTTACGGATACTATGGAGCTTGAATTGGTTACGAATTGAAAGAGCCTA
 AGTATACGCTTGAAGAAGCTCGTATCCACGATGCATCTTATTCAAGCACCTATTTGTAACTT
 40 CCGTTGATTAATAAAGAAACAGGAGAAATCAAACACTCAAGAAGTTCTCGGTGATTTCCA
 ATTATGACTGAAATGGGTACATTCACTCAACGGTGGTGAACGTATTATCGTTCTCAGTTGG
 TTGTTCTCCTGGTGTATTCAACGATAAAAGTTGATAAAAACGGTAAAGTTGGTTACGGTTC
 AACTGTAATCCCTAACCGTGGAGCATGGCTTGAATTAGAAACAGATTCAAAGATATTGCTAC
 ACACGTATCGACCGTACACGTTAAATTCAACTCTTGTACGTGCGCTTGGTTCTCAG
 45 GTGATGATGAAATCATGGATATCTTGGTGAAGCGAACTTGTCTGTAACACAATCGAAAAGA
 TATTCAACAAAACCCAGCAGACTCACGTACTGACCGAAGCTCTAAAGAAATTACGAACGCCTT
 CGTCCAGGTGAACCAAAACAGCTGATAGCTCACGTAGCTTGTAGCTCGTTCTTGTGACC
 CACGTCGTTATGACTTGGCAGCTGGTGTACAAAATCAACAAAAACTAACATCAAGAC
 TCGTCTTGTGAAACAAACATCGCTGAAACTTGGTTGATGCTGAAACTGGTGAATCCTTGT
 50 GAAAGCTGGTACAGTAATGACACGTGACGTGATTGATTCAATCGTGTACATGGATGGTGA
 TTAACAAATTGTTACACACAAATGATTACGCTGTTGTCAGTGAACCTGTTCTTCAAAA
 ATTCAAAGTTGTTGCACCAAACGATCCACAGACCGCGFTGTTACAATCGTTGGTAACGCAAATCCT
 GATGACAAAGCGCGTGCCTACACCAGCTGATATCTTGGCAGAAATGTCTTACTTCTTAAACC
 TTGCTGAAGGTCTAGGTAAGTTGATGATATCGACCACTTGGGATCGTGTATTGCGCCGT
 55 TGGTGAATTGCTGCTAACCAATTCCGTATTGGTCTTGCTCGTATGGAACGTAACGTTGGGAA
 CGTATGTCAGTTCAAGACAACGAAGTGGACACCACAAACATCAACATTGTCCTGTTA

CTGCAGCCGTTAAAGAACATTCTCGGTCATCTCAATTGTACAGTCATGGACCAACACAACCC
 ACTTTCTGAGTTGTCTCACAAACGTCGTTGTCAGCCTTAGGACCTGGTGGTTGACTCGTGAC
 CGTGTGGTTATGAAGTGTGACGTGCACTACACTCACTATGGTCGTATGTGTCCGATTGAAA
 CTCCCTGAAGGACCTAACATCGGTTGATCAATAACTTGTCAACATACGGACACCTAACATAAATA
 5 TGGTTTCATCCAAACACCATACTGTAAGTTGACCGCGCTACAGGTGTGATTACAAACGAAATC
 GTTGGTTGACTGCCGATGAAGAAGATGAATAACACAGTAGCACAGGCTAACTCAAAACTTAACG
 AAGATGGAACATTTGCTGAAGACATCGTTATGGGACGTACCAAGGTAAATAACCAAGAGTTCCC
 AGCAAGCGTTGTGACTTCGTAGACGTTCACCTAAACAAGTAGTTGCCGTTGCGACAGCATGT
 ATTCCCTTCCTGAAAACGATGACTCTAACCGTGCCCTTATGGGTGCCAACATGCAACGTCAAG
 10 CGGTGCCATTGATTGATCCACACGCACCATACTGTTGGTACTGGTATGGAATATCAAGCAGCCA
 CGACTCAGGTGCTGCAGTTATCGCTAAACACGATGGACGCGTTATCTTCTGATGCTGAAAAAA
 GTTGAAGTTCGTCGCGAAGATGGTTCACTTGATGTTACACATTACTAAATTCCGTCGTTCTA
 ACTCAGGTACAGCTTATAACCAACATACACTTGTAAAGTTGGCGATATCGTTGAAAAAGGTGA
 CTTCATCGCTGATGGTCCTCAATGGAAAAAGGTGAAATGGCCTTGGTCAAAACCCAATCGTC
 15 GCTTACATGACTTGGGATGGTTATAACTATGAAGATGCCATCATCTTGAGTGAACGTCTTGTAA
 AAGAAGATGTTATACATCAGTTCACTTGGAAAGAATTGAAATCAGAAACACGTGATACTAAGTT STRF
 AGGCCCTGAAGAAATCACTCGCAGATTCCAAACGTTGGTGAAGAAGCTCTAAAGACCTTGAC
 GAAATGGGTATTATCCGTATCGGTGCTGAAGTTAAAGAAGGTGACATCCTGTAGGTAAAGTAA
 CACCTAAAGGTGAAAAAGACCTTCTGCTGAAGAGCGCCTCTCACGCAATCTTGGTGTATAA
 20 ATCACGTGAAGTCGTGATACATCACTTCGTGTACCCACACGGTGGAGATGGTGTGTTGAC
 GTTAAAATCTTACACGTGCAAACCGTGATGAATTACAATCAGGTGTTAACATGCTGTTGAC
 TTTATATCGCACAAAACGTTAAAGTCAAGTCGGAGATAAAATGGCCGGTCGTCACGGTAACAA
 AGGGGTTTCTCGTGTGTTCCAGTTGAAGACATGCCCTATCTTCCAGACGGAACCTCAGTC
 GATATCATGTTGAACCCACTTGGGGTGCCTCTCGTATGAAACATCGGACAAGTTATGGAGCTTC
 25 ACCTTGGTATGGCTGCTGTAACCTTGGTATTCACATTGCAACACCAGTCTTGTATGGGCAAC
 TTCTGAAGACCTTGGGATACAGTTAACGAAGCTGGTATGGCTAGCGACGCTAACAGACAGTTCTT
 TACGATGGACGTACTGGTGAACCATTGATAACCGTGTGTCAGTTGGTGTATGTACATGATTA
 AACTTCACCACATGGTTGATGATAAAACTCACGACGTTCAAGTTGGCTTACTCACTTGTAC STRR
 30 GCAACAAACCTTGGTAAAGCACAATTGGTGGACAAACGTTGGTGAATGGAAAGTTGG
 GCTTGGGAAGCTTACGGTCATCAAATGTTCTCAAGAAATCTTGACTTACAAATCAGATGATG
 TCAACGGTGTCTTAAAGCTTATGAAGCCATCACTAAAGTAAACCAATTCCAAAACCAGGTGT
 TCCAGAAATCATTCCGAGTTCTGTTAAAGAATTGCAATCACTTGGTCTTGACATGCCGTGCTT
 GATGAAGATGACAATGAAGTAGAACCTCGTGATCTGATGAAGGTGAAGATGACGATGTTATGC
 35 ACGTTGATGATCTGAAAAAGCTCGTCAAAACAAAGAAGCAGAAGAAGCGGAAAAAGCAGAAGT
 TTCTGCAGAAGAAAACAAATAATAGGAAAGAACATTCAAGACATGAGAGAGGCAAGACCTGCTTC
 TCTTGGTCAGATTGTTGATTGAGTCCTATAACGATAAAATGATGTCTTACGAATCATGAATTG
 TAAGTCATGACAGTTAGAAAGTAGCGCAGCTATTCAAAGTCATAAGAAGGTATCATGGTACG
 TAATCGTTACAGCCGGCGTC -3'

40 SEQ ID N°3 : Séquence du gène *rpoB* d'*Abiotrophia defectiva*. Cette séquence mesure 4 325 paires de bases, possède un contenu en cytosine plus guanosine de 47 %, et est déposée dans GenBank sous le numéro AF 535173 :

5' -ATATAGGGCACCGCGTGGTCGACGGCCGGCTGGTCCTAAACACATGTAACGTCACTCCGATG
 AGTTGGTTCTGTTGCTTTTTGCGCTTCAAAGACGAAAAATGTCATTGTCAACAATTAT
 45 TAATAATTGTAACCTTAATGTAAGTGGTGTCTTAGATTATATTAGGGTGAATCGCTTGA
 GTCATATCGTAAACACGGTAAAAAGCTGAGCGTCGAAGCTATGCCGTATCGACGAAGTCTT
 AGAGTTGCCGAACCTGATTGAAATCAAACGGATTCTACAAATGGTCTTGGATGAAGGGCTA
 AAAGTGATGTTGAGGGACATTGCGCATTGTCGACCAATTGGAGAACCTGGAAACTTCATTTG
 TAGACTATGAGTTCAAGGAAGCTAAGTATAGCTTAGAAGACGCTGTAGCCATGACGCTAACTA
 50 CTCAAAACCAATCTATGTAACCTGCGCCTGTTCAACAAAGAGACAGGTGAAGTCAAAGAACAA
 GAAGTCTCTCGGGGACTTCCCAATCATGACCGAAATGGGACCTTCATTATCAACGGGGCG
 AACGGGTTATCGTTCCAGTTGGTACGTTCTCCAGGTGTCTACTTCCACGACCGTATGGACAA
 GAAAGGCCACAGCTATAACTCTACGGTTATTCTAACCGTGGGGCTTGGTGAATTGAA
 TCAGATGCTAAGGGATTGCCTACGTCCGCATTGACCGGACCCGGAAAGATTCCATTGACTGTCT
 55 TGATGCGTGCCTTAGGTTGGTCAAGATGACGAGATTATGATATCTCGGCCAATCTGAGCT
 CTTAGACTTAACATCGAGAAGGATGTTCACAAAAACATTCAAGACTCTCGTACGGAAGAAC
 TTGAAGGACATTACGAGCGTCTCCGTCCAGGTGAACCTAACGACCGCAGAAAGCTCACGTAACC

TCTTGGTTGCCGCTTCTGACCCACGTCGCTATGACTTAGCACCTGTAGGTCGTATAAGAT
 CAATAAAAAGCTCCACCTCAAGAACCGTTGGTGGCTTGACTTTGGCTGAAACCTTGGTTAAC
 CCAGAAACAGGCGAAGTGCTCTTGAAAGAAGGAACGGCTTGGAAGAACGTGTTCAAGCCC
 TGATTCCATACTTAGAGGCTGGCTTGAAATAAGGTAAACCCCTATCCTCTGAAGATAGTGTGGT
 5 AGCTCAACCAATTGATTACAAATCATCAAAGTTATTCACCTAAGAACGCCAGCAAGTGATT
 AACATCATCGGTAACGGGAACATTGAGAAGATTAAGTGCTTGACGCCAGCTGACATTATTGCGT
 CAATGAACACTATCTCTATTAGACCAAGGAATTGGTGTGACAGATGATATCGACCACTGGC
 TAACCGTCGTATTCGTTAGTCGGTGAATTATTGAAAACCAATTCCGTATCGGGCTATCCC
 10 ATGGAACGGGTAGTGCCTGAACGTATGTCGCTCCAAGATGTTGCGACCACACCGCAACAAT
 TGATTAACATTCGTCAGTAGTGGCGCTATTAAGGAATTCTTCGGTCATCCCAGTTGTCACA
 ATTCAATGGACCAAGTTAACCCACTCGGGGATTGACCCACAAACGTCGTGTCAGCCTTAGGG
 CCTGGTGGTTGACGCCGGACCGTGCCGGCTATGAAGTGCGGGACGTTCACTACTCTCACTACG
 GCCGTATGTTGTCACGAGACGCCAGAAGGTCTAACATCGGGTTGATTAACAGCTTGTCTTC
 15 TTATGCCAAGATTAACAAGTATGGTTATTGAGACGCCTTACCGTAAAGTGGACAAATCGGTT
 ACGCCACACCGTGTACGACCGAAATTGACTACCTAGCAGCGGACGAGGAAGACTGTACGTAG
 TAGCCCAAGCCAACCTCTAAACTCAACGAAGACGGGACCTCGCCAATGACCTAGTTATGGCGC
 TTTCCGTTACAAAACATTGAGGTTAACGTTGACCAAGTAGACTACATGGACGTATGCCAAAA
 CAGGTTGTCGCTGCGCACTGCTAGCATTCCGTTCTGGAAAACGACGACTCCAACCGGGCT
 20 TGATGGGTGCCAACATGCAACGTCAAGCTGTGCCACTTATTAATCCACAATCCCCACTGATTGG
 GACTGGGATGGAATATAAGGCAGCACACGACTCTGGGGCTGCGCTTATGTAAGCGCCCGT
 GAAGTGGTTATGCGATGCTAACAAAGGTGCGGTGCGCACTCCAGAAGGTGAAGTTGACCAAT
 ACCGTTAACCAAGTTGACGTTCTAACGCTGGGACCTGTTACAAACCAACGTCCAATCGTAGA
 ATTAGGCGACCAAGTTGATGCCTTGGAAATCTTAGCAGATGGTCATCTATGCAAAATGGGAG
 25 ATGGCCCTCGGTCAAAACCACTGGTAGCCTCATGACTTGGGAAGGGTATAACTATGAGGACG
 CGGTTATCATGTCACGCTGGTCAAAGACGATGTTATACCTCATCCACATTGAAGAATA
 TGAATCAGAGTCCCGTGYACYAAGTTAGGCCCTGAAGAAATTACACGCGAAATTCCAACGTG STRF
 TCCGAAGATGCCCTCAAGTACTTAGACAAAGACGGGATTATCTGATCGGGCGGAAGTAAAG
 ACGCGATATCTTAGTTGGTAAGGTAAACACCAAAAGGTGTGACCGAGTTGTCTGGGAAGAACG
 CTTGCTCCATGCTATCTCGGTGAGAAGCGCGTGAAGTACGTGATACTTCCCTGCGTGTGCCA
 30 CACGGGGGGCGGGATTGTCACGCTTAAATCTTACCCGCGAAGCTGGCACGAATTGG
 CACCAAGGTGTCAACAAGCTAGTCGCGTCTACATCGTACAAAACGTAAAATCAATGAAGGGGA
 TAAGATGGCCGGTGTACGGTAACAAAGGGTTGTCTCCATTACATGCCGAAGAAGATATG
 CCATTCTTACCAAGATGGTACCCAGTTGATATCATGTTGAACCCATTAGGGTTCCATCCCGTA
 TGAACATCGGGCAAGTCTAGAGTTACACTTGGGATGGCTGTCGCAAAATGGCATCAAGAT
 35 TGCAACACCTGCTTGTACGGCTAGTGAAGAAGATGTCCTGGAAACAGTTAAGGAAGCCGGC
 TTAGAAGCTGACGCTAACAGACTATCTTATATGATGGTCGAACCGGTGAACCATTGACCGTAAAG
 TCTCTGTTGGGTTATGTACATGATTAAGTTGGCCACATGGTCGATGACAAGTTGACGCCCG STRR
 TTCAACAGGTCCATACTCTGGTTACCAACAAACATTGGTGTAAAGCTCAATTGGTGGG
 CAACGTTCGGGGAGATGGAGGTTGGCCCTA -3'
 40

SEQ ID N°4 : Séquence partielle du gène *rpoB* de *Streptococcus mutans*. Cette séquence mesure 3198 paires de bases, elle possède un contenu en cytosine plus guanosine de 42 %, et est déposée dans GenBank sous le numéro AF 535167.

45 5' -GGACCCTTTATGACTTCTGGATACAGGTCTGAAGGAAGTTTTGAAGATGTGCTTCCAATT
 CCAATTTCACAGACACTATGGAATTAGAGTTGTGGTTATGAGTTGAAAGAGCCTAACGTATAC
 ATTGGAAAGAACGTGCTCATGATGCACATTATTCTGCCCTATCTTGTACTTCCGTCTC
 ATCAATAAGAAACTGGTGAATTAAAGACACAAGAAGTATTTTGTTGGTGAATTCCCTTGATGA
 CTGAAATGGGTACTTTATTATTAATGGTGTGAACGTATTATCGTTCTCAGTTGGTACGTT
 50 ACCAGGTTTATTTAATGATAAAAGTGGATAAAAATGGAAAATTGGCTATGGTCAACTGTT
 ATCCCTAACCGCGGTGCTTGGCTTGAGCTTGAAACGGACTCTAACGGATATTGCTTATACTCGTA
 TTGATCGTACTCGTAAATTCCCTTTACGACGCTGGTGTGCACTCGGTTTCCGGGGATGA
 TGAGATTATTGATATTTGGTGTAGCGAATTGGTGTGAAATACCAATTGAAAAGATATCCAT
 AAAAATCTAACGACTCTCGTACAGATGAAGCTCTCAAGGAANTTATGAACGTCTCGTCCGGG
 55 TGAACCTAAAACGGCAGATTNCNTCACGCAGTCTTGATTGCACGTTTGATGCGCGCCGT
 TATGATTAGCAGCTGTTGGCGCTATAGATAATAAGAAGTTAACGTAAAACGGTCTTGAA

TCAAGTCATTGGCTGAAAANNAGATCTGAAACAGGCAGAAATTCTTGTGAAAGCTGGACT
 GAAATGACACGCAGTGTATTGATTGATTCGAGATTATCTTGATGGAGATCTCAATAAAATTG
 TTTATACGCCAAATGAATACGCTGTTGACAGAACCTGTTCTTCAAAAATTCAAAGTTAT
 GGCTCCAAATGATCCAGACCGCACGGTTACTGTTATTGGTAATGCCAGTCCAAGATGACAAAGT
 5 ACGTCACCTGACACCAGCCGATACGTATTAGCTGAAATGTCTTATTCTCTTAACGGCTGAGG
 GTNTAGGTAAAGTTGATGATATTGACCATTAGGCAACCGACGTATCGTGTGTTGGTGAATT
 GCTTGCTAATCAATTCTGATTGGTTGGCACGTATGGAACGCAATGTTGCTGAACGCATGTCC
 GTTCAAGATAATGAAGTCTTAACGCCACAAACAGATTATTAACATTGCCCTGTAACAGCGGCAA
 10 TTAAAGAGTTTTGGTTCTTCTCAATTGTCACAGTTCATGGACCAACACAATCCACTGTCTGA
 ATTGTCTCATAAACGCCGTTGTCAGCTTAGGTCCTGGTTAACACCGACCGTGCTGGT
 TATGAAGTCCGTATGTGCACATACGCATTATGGTCGTATGTCCAATTGAAACGCCCTGAAG
 GACCAAATATTGGATTGATTAATAACTTGTCTTCTATGGTCATCTTAATAAAATATGGATTAT
 CCAAACACCATACCGTAAAGTTGACCGTGAGACAGGTAAGTAACCAATGAAATGAAATGGCTT
 15 ACTGCTGATGAAGAAGATGAATTCACTGTAGCTCAGGCTAACTCAAAACTCAATGAAGATGGAA STRF
GCTTTGCTGAAGAAATCGTCATGGGACGTCATCAAGGAAATAACCAAGAGTTCCAGCAAGTTC
 TGTTGAATATATGGATTTCTCTTAAGCAGGTAGTTGGCGTAGGACAGCATGTATTCTTTC
 CTTGAAAATGATGACTCCAACCGTGCCTTATGGGAGCTAACATGCAGGCCAAGCTGTGCCAT
 TGATTGATCCTAAAGCACCTTTGGAACTGGTATGGAATATCAAGCAGCCATGATTCTGG
 AGCCGCTATTATCGCTAACATAATGGGAAAGTGGTTATTCCGATGCAAGATAAGATTGAAGTT
 20 CGCCGTGAAGATGGCTCACTAGATGTTATCATGTTACCAAATTCCGTCCTTAACCTGGAA
 CTGCCTACAATCAACGTACTCTGTTAGGGTAGGCGATAGTGTGAGAAGGGGGACTTTATTGC
 AGATGGTCTTCATGGAAAAGGGTGAGATGGCTCTGGACAAAATCCAGTGGTTGCTTACATG
 ACTTGGGAGGGTTACAACCTTGAAGATGCTGTATCATGAGCGAGCGTCTGTCAAGGATGATG
 TTTATACCTCTGCCATTAGAAGAATTGAATCTGAAACTCGTACAAAGCTGGACCTGA
 25 AGAAATTACCGTGAAATCCAAATGTTGGTGAAGATGCCCTGAAAGACCTTGATGAAATGGGA
 ATTATTCCGATTGGTGCTGAGGTAAAGAAGGTGATATTCTAGTTGGTAAAGTGACTCCTAAAG
 GAGAAAAGATCTTCTGCAGAACGCGCTTGCATGCCATTGGTGAACAAATCACGTGA
 AGTTCGTACACTCTCGTGTACCTCATGGTGGCGACGGTGTGTTGTGATGTGAAATC
 TTTACACGTGCTAATGGAGATGAACCTCAATCAGGTGTTAACATGCTGGTTCGTGTTATATCG
 30 CTCAAAACGTAATCAAGGTGGAGATAAGATGGCGGACGTACGGTAACAAAGGGTGTGCT
 TTCCCGTATTGTACCGAGTGGAAAGATGCCATATCTCCAGATGGAACACCTGTTGATATCATG
 CTTAACACTGGGGTGCCATACGGATGAAACATTGGCAAGTTATGGAACCTCCATCTGGTA
 TGGCTGCTCGTAATTGGCATTGATATTGCAACGCCCTGTCTTGACGGAGCAACTCTGATGA
 TCTTGGAAACAGTAAAAGAACGGGTATGGATTCTGATGCTAAACTGTTCTTATGATGGT
 35 CGCACAGGGAGCCGTTGATAATCGTGTATCAGTTGGTGTATGATGAAACTTCACC STRR
ACATGGTTGATGAYAACCATTTGTCTATGCAAGWTCAGTTGGCCCTAKTCAYGAWTAMTC
 AGASGARTTCCTGCTWGGTAAAGGCTNCAATTGCTTTAGAGGTTAAGGCTGGTGAATAAC
 GGTATGCTGGTATTGATGGCAATGGCAAGTGAATANTCAACACCGGGCGTACANCGTGC-3'

40 SEQ ID N°5 : Séquence partielle du gène *rpoB* d'*Enterococcus faecalis*. Cette
 séquence mesure 3096 paires de bases, elle possède un contenu en cytosine
 plus guanosine de 42 %, et est déposée dans GenBank sous le numéro AF
 535175

45 5'-GACCCTTATCAATTGGTTTAGATGAGGGACTTCGTGAAATGTTGAAGACATTTCACCAATT
 GATGATTCCAAGGAAACTTATCCTTAGAATTGTTGACTATGAATTAAAAGAACCAAAGTACA
 CAGTAGAAGAACGCCGCGCACATGATGCCAACTATTCTGCCATTACATGTAACATTACGTT
 AACCAACCGTAAACAGGTGAAATTAAATCCAAAGAAGTCTTCTCGCGATTCCATTAAATG
 ACAGAAATGGGTACCTCATCAACGGGGCAGAACGTGTTATCGTTCCAATTAGTCGTT
 CTCCAGGTGTTACTCCATGGAAAAGTGGACAAAAACGGCAAAGAAGGTTGGCTCAACAGT
 50 CATTCTAACCGTGGTGCATGGTTAGAAATGGAAACAGATGCGAAAGACATTCTTATGTTCGG
 ATTGACCGCACACGTAAAATTCTTAACTGTTAGTTCGTGTCTTAGGTTCGGTTAGATG
 ATACCATCTCGAAATTTCGGCGACAGCGAAAGCTACGCAACACAATTGAAAAAGATTACA
 CAAAAACGCAAGTGTACAGAACAGGCTTGAAAGACATTATGAAACGTCTCGCCCA
 GGCAGAACCAAAACAGCAGATAGCTCACGTAGCTTGTAACTTGCACGTTCTTGATCCAAA
 55 CGTTATGATTGGCAAACGTTGGCGTACAAAGTTAACAAAAATTAGACTTAAACACAGTC
 TATTAAACTAACCTTAGCTGAAACGCTAGTTGATCCAGAAACTGGTGTAAATCATTGTCGAAA

AAGGCACAGTTAACACACTACATGGAAACATTAAGGCRATACATTGACAAACGGCTAA
 ACAGCGTAACTTACTATCCAAGTGAAGATGCGGTAGTAACGTAACTGAACCAATGACGATCCAAGTGAT
 TCAAGTTCTTCAACAAAAGATCCTGAACGTATCGTAAATGTGATTGGTAACGGCTATCCAGAC
 GACAGCGTAAAAACAGTCGTCCAGCAGATATCGTTGCTTCAATGAGCTACTTCTCAACTTAA
 5 TGGAAAGATATCGGTAATGTCGATGACATCGACCACTTAGGTAATCGTCGTATCCGTTAGTAGG
 CGAATTATTACAAAACCAATTCCGTATTGGTTAGCCCCTATGGAACGTGTGGTTCTGTGAAAGA
 ATGTCTATTCAAGACACAGAACATTGACACCAACAAATTAAACATCCGTCCAGTGGTAG
 CAAGTATCAAAGAATTCTTGGTTCTCACAGTTATCAGTCAGTTCATGGACCAAACAAACCCATT
 AGGTGAGTTAACCCATAAACGTCGTCTATCAGCCTAGGGCCTGGTGGTTGACTCGTGTATCGT
 10 GCCGGTTATGAAGTTCGTACGTTCACTACTCTCACTATGGTCGTATGTGTCCAATTGAAACGC
 CTGAGGGACCAAATATCGGGTTGATCAATAGCTTATCTAGTTATGCGAAAGTGAATTAAATTGG
 TTTCATCGAAACGCCCTATGCCGTGTTGATCGTGCACAGGCCGTGTTACTGATCAAGTAGAT
 TACTTAACAGCAGACATCGAAGACCAATTATCGTAGCGCAAGCGAACTCACTTTAAATGAAG
 ATGGCACATTGCCAATGATGTTATGGCGCGTCTACAAAGTAAAAGCTAGAAGTTGCCGT
 15 AGACAAAGTTGACTACATGGACGTTACCAAAACAAGTAGTCGAGTCGAAACAGCATGTATT
 CCTTTCTTAGAAAACGATGACTCCAACCGTGCCTGATGGGTGCCAACATGCGAGCGTCAAGCGG
 TGCCGTTAATTCAACCACGCTCTCCGTGGTAGGTACAGGTATGAAATATAATCAGCCATGA
 CTCAGGTGCTGCTTACTATGTAACATGACGGTGTGCTAGAATTGTCGATGCAAAAGAAATT STRF
 CGCGTTCGTCGCGACAATGGCGATTAGACAAATATATGGTTACAAAATTCCGTGTTCTA
 20 CAGGAACAAGCTACAACCAACGCCAATTGTCACTTAGGTGAAAAGTTGAAAAGGCATACTT
 TACCGGATGGACCTCTATGGAAGAAGCGAAATGGCTTATGGCAAAACGCTTATGCTTC
 ATGACATGGGAAGGTTACAACCTACGAGGATGCCATTATCATGAGCCGTCGTTAGTTAAAGACG
 ATGTCTACACTCTGTGCATATTGAAGAATATGAATCAGAACGTGATACAAAATTAGGACC
 TGAAGAAATTACCGTGAAATTCCAAACGTTGGGGAGACGCGTTGAAAGACTTAGACGAAATG
 25 GGGATTATCCGCAATTGGTGTGAGTTCAAGATGGCGACTTACTAGTTGGAAAGTCACACCTA
 AAGGGGTACAGAATTATCTGCAGAAGAACGTTATTACACGCAATCTCGGGAAAAAGCCCG
 CGAAGTTCGTGTACGTCCTCCGTGTACCTCACGGTGGCGCGTATCGTTATGATGTGAAA
 ATCTTACTCGTGAAGCTGGCGATGAATTATCACCAAGGTGTCAACATGTTAGTTCGTGTCTATA
 TCGTCAAAACGTAATTACCGAAGGGAGATAAAATGGCGGGACGTCACGGAAATAAAGGGGT
 30 TGTTTCCCGTATTATGCCGGAAGAAGATATGCCATTCTACCTGACGGAACACCTGTTGATATC
 ATGTTGAACCCATTAGGGTACCTTCTCGTATGAATATCGGACAAGTACTTGAATTACACTTAG
 GTATGGCTGCTGCCAATTAGGTATTACCGTCAACACCTGTTGATGGGCAACCGATGAG
 AGACGTTGGGAAACTGTTCGTGAAGCTGGTATGGCTAGCGATGCTAAAACAGTTCTTACGAT
 GGACGTACAGGTGAACCAATTGATAACCGTATTCCGTTGGTGTATGTTGATATTAAATTAG
 35 CCCACATGGTTGATGACAAATTGCATGCTCGTTCAATCGGACCTACTCTCTGTTACGCAACA STRR
 ACCGTTGGGTGTAAAGCTCAATTTC-3'

Dans les séquences qui précèdent, le nucléotide K désigne T ou G, le nucléotide M désigne A ou C, le nucléotide R désigne A ou G, le nucléotide W désigne A ou T, le nucléotide Y désigne C ou T et le nucléotide N désigne A, T, C ou G.

Exemple 2 : Séquençage partiel du gène *rpoB* de 28 espèces du genre *Streptococcus* et genres apparentés.

A partir de l'alignement des séquences complètes du gènes *rpoB* chez *Streptococcus* spp. et *Abiotrophia defectiva* de l'exemple 1 et celles connues dans GenBank, (*Streptococcus pneumoniae* AE008542 et *Streptococcus pyogenes* AE006480) un jeu d'amorces a été choisi pour l'amplification et le séquençage d'un fragment 709 à 740 pb de ce gène chez 28 souches type de ces genres bactériennes. Ces amorces ont pour séquence :

- SEQ ID N°6: 5'-AARYTIGMCCTGAAGAAAT-3'
- SEQ ID N°7: 5'-TGIARTTIRTCATCAACCATGTG-3'

La séquence SEQ ID n° 7 est utilisée à titre d'amorce 3' et représente donc la séquence inverse complémentaire du brin direct représenté dans les 5 séquences SEQ ID n° 1 à 5 qui précèdent.

... Ces amores sont incorporées avec l'ADN extrait des bactéries dans une PCR selon les conditions suivantes : dénaturation à 95°C pendant 1 min suivie de 35 cycles comportant une étape de dénaturation à 94°C pendant 10 sec, une étape d'hybridation à 52°C pendant 10 sec et une étape d'elongation à 72°C 10 pendant 30 sec.

Les produits amplifiés sont séquencés par les mêmes amores SEQ ID N°6 et SEQ ID N°7 selon les conditions suivantes : dénaturation à 95°C pendant 1 min suivie de 30 cycles comportant une étape de dénaturation à 95°C pendant 30 sec, une étape d'hybridation à 52°C pendant 30 sec et une étape 15 d'hybridation à 62°C pendant 1 min. Les produits de séquençage sont analysés par un séquenceur ABI PRISM 3100.

Les inventeurs ont déterminé la position de ces deux amores SEQ.ID. n° 6 et SEQ.ID. n° 7, de façon à respecter les critères suivants :

- 1- séquence encadrée par ces deux amores spécifiques de l'espèce de la bactérie. Cette condition est vérifiée après alignement des fragments de environ 20 720 pb avec l'ensemble des séquences des gènes bactériens *rpoB* disponibles dans les banques de données informatiques.
- 2- recherche d'une région d'identification la plus courte possible afin d'augmenter le plus possible la sensibilité de la détection moléculaire
- 25 3- la longueur des amores de 18 à 22 pb,
- 4- séquence des amores présentant une température de fusion voisine,
- 5- séquence des amores ne permettant pas d'auto-hybridation ni de complémentarité.

Les fragments obtenus du gène *rpoB* des bactéries des espèces du genre 30 *Streptococcus* et desdits genres apparentés ont environ 720 (709 A 732) paires de bases et leur séquence est spécifique de chaque espèce de ce genre et permettant donc l'identification moléculaire des bactéries des 28 espèces testées sont :

SEQ.ID. n°8 : séquence partielle du gène *rpoB* chez *Streptococcus suis* CIP 1032 17^T mesurant 709 paires de bases :

5' – CGCGAAATTCCAAACGTTGGTGAAGATGCCCTCGCAACTGGACGAAA
 5 TGGGGATTATCCGTATTGGTGCGAAGTTAAAGAGGGCGACATTCTTGTGG
 . TAAAGTCACACCAAAAGGTGAAAAAGATCTTCTGCTGAAGAGCGTCTCTGC
 ACGCAATCTCGGTGACAAGTCACGTGAAGTACGTGATAACCTCTCTCGTGTGTA
 CCTCACGGTGCCGATGGTGTCTCGTGTGAAATCTTACTCGTGCCAA
 CGGTGATGAATTGCAATCAGGTGTTAACATGTTGGTCTGTGTTACATCGCTC
 10 AAAAACGTAAGATCAAGGTGGAGATAAGATGGCCGGTCGTACGGTAACAA
 GGGTGTCTTACGTATTGTACCTGTTGAGGATATGCCATATCTCCAGATG
 GAACACCAGTTGACATCATGTTGAACCCACTCGGGTGCCATCACGTATGAAC
 ATCGGTCAAGGTTATGGAACCTCACTGGGTATGGCGGCTCGCAACTGGGCA
 TCCATATCGCAACACCAGTTTCGATGGTGCAAGTTAGAAGACCTCTGGTCA
 15 ACTGTTAAAGAACCGAGGTATGGACTCAGATGCCAAGACCATTCTTACGATGG
 ACGTACAGGTGAACCATTGACAACCGTGTATCTGTTGGTGTATGTACATGA
 TCAAGCTTCACCACATGGTTGATGACA – 3'

SEQ.ID.n°9 : séquence partielle du gène *rpoB* *Streptococcus sanguinis* CIP

20 55.128^T mesurant 725 paires de bases :

5'- TGTCAACCATGTGGTGAGCTTAATCATGTACATGACACCGACAGATA
 CACGGTTGTCAAACGGCTCACCGGTACGTCCATCGTAAAGAACAGTCTGGCA
 TCGCTATCCATACCAAGCTTCACGGACAGTATCCCAGAGGTCTCTGAGCTG
 TCCATCAAAGACCGGTGTCGCAATATGGATGCCAAGTTACGTGCTGCCATAC
 25 CAAGGTGAAGCTCCATAACCTGACCAATGTTACGTGATGGTACCCGAGT
 GGGTTAGCATGATATCAACTGGTGTCCGTCTGGCAAATAAGGCATGTCTTC
 CACAGGAACGATACGGGATACAACCCCTTGTCTCCGTGACGACCAGCCATCT
 TATCTCCGACCTTGATCTACGTTTGAGCGATGTAGACACGAACCAACATAT
 TAACGCCAGATTGCAACTCATCACCATTAGCACGGTAAAGATCTCACGTCA
 30 CGAACCACTCCATCAGCACCGTGGCACACGCAGAGAGGGTATCACGGACTTC
 ACGAGACTTGTCTCCGAAGATAGCGTGCAAGAGGGCGCTTCAGCAGAAAGA
 TCTTTTCACCCCTTAGGGTAACCTTACCTACAAGGATATGCCCTCCTGACT
 TCCGCCCGATGCGGATAATACCCATTGTCGAAATTGCGTAGGGCATCTTC
 CCCTACGTTGGAATTGCGGGTAATTCTTCAGGTCA – 3'

SEQ.ID. n°10 : séquence partielle du gène *rpoB* *Streptococcus salivarius* CIP 102503^T mesurant 728 paires de bases :

5'- TTGTCATCAACCATGTGTGAAGTTGATCATGTACATGACACCAACTGAT
5 ACACGGTTATCAAATGGTCACCTGTACGTCCATCGTAAAGGATTGTCTTAGC
ATCACTATCCATACCTGCTCACGAACAGTATCCCAGAGGTCTCTGAGCTTGC
CCCGTCAAAGACTGGTGTGCGATGTGGATACCAAGTTACGAGCAGCCATA
CCAAGGTGAAGTCCATAACCTGACCGATGTTACACGTGATGGCACCCCAAG
AGGGTTAACATGATATCAACTGGTGTACCGTCTGGAAGGTAAGGCATGTCT
10 TCAACACAGGAACAATACGAGAAACAACCCCTTGTACCGTGACGACCGGCCAT
CTTATCTCCGACCTTAATCTTACGTTTGTGAGCGATGTAACACGAACAAAGCAT
GTTAACACCTGATTGCAATTCATCACCCTTGCACGTGTGAAGATTAAACATC
ACGAACGACACCATTACCAACCCTGAGGTACACGGAGTGAGGTATACGTACT
TCACGAGATTATCACCAAAGATAGCATGGAGAAGACGTTCTCAGCAGAAA
15 GGTCTTTTCACCCCTAGGTGTTACCTTACCAACAAGAATGTCACCTTCTTAA
CCTCAGCACCGATAACGGATAATACCCATTCTGTCAAGGTCTTGAGAGCTTCTT
CACCAACGTTGGCAATTACGTGTAATTCTTCTCAGGTCCA – 3'

SEQ.ID. n°11 : séquence partielle du gène *rpoB* chez *Streptococcus pyogenes*
20 CIP 56.41^T mesurant 725 paires de bases :

5'- TGTCAACCATGTGGTGAAGTTGATCATATACATGACACCAACGGAT
ACACGGTTGTCAAATGGTCACCGGTGCGACCATCATAAAGGACCGTCTTAGC
ATCGCTATCCATACCAGCTCACGAACAGTGTCCAAAGGTCTCTGATGAAG
CCCCGTCAAAGACAGGTGTTGCAATGTGAATACCAAGATTACGAGCAGCCATA
25 CCAAGGTGAAGTCCATAACCTGACCAATATTACCGTGATGGCACCCCAAG
AGGGTTAACATGATGTCAACTGGTGTCCGTCTGGAAGGTATGGCATGTCT
TCAACTGGTACAATACGTGAAACGACACCCCTTGTGACGACCGGCCAT
TTTATCTCCGACCTTGATTTCACGTTTGTGAGCGATGTAACACGCACAAGCAT
ATTAACACCTGATTGCAATTACGCCGTAGCGCGTGTAAAGATTTCACATC
30 ACGAACGATAACCATTACCAACCCTGAGGGACACGAAGTGAGGTATACGCAC
TCACGCGATTATCCCCAAAGATGGCGTGAAGTAAACGTTCTCAGCAGAAAG
GTCCTTTTACCTTAGGTGTGACTTTACCTACTAAGATGTGCCCTCTTAAAC
CTCAGCACCGATAACGGATAATGCCATTCTGTCAAGGTCTTGAGGGCTTCTT
CACCAACATTGGGATTCCGAGTGATTCTCAGGGCA – 3'

SEQ.ID. n°12 : séquence partielle du gène *rpoB* chez *Streptococcus pneumoniae* CIP 102911^T mesurant 724 paires de bases :

5' – CAACCATGTGGTGGAGTTGATCATGTACATGACTCCGACAGAAAACACG
 5 GTTATCAAACGGTCACCAAGTACGTCCATCGTAAAGGATCGTTGGCATCGC
 TATCCATACTGCTTCTTAACAGTTGACCAAAAGATCTCAGAACTGCTCCAT
 CAAAGACTGGTGTGCGATGTGAATACCAAGAGTACGAGCTGCCATACCAAG
 GTGAAGCTCCATAACCTGACCGATATTACATACGTGATGGTACCCCAAGTGGGT
 TCAACATGATGTCGACTGGAGTTCCGTCTGGAAGGTAAGGCATGTCTTACA
 10 GGAACGATAACGAGAGACAACCCTTGTTCCGTGACGTCCGGCCATTATC
 TCCGACCTTAATCTTACGTTTGAGCGATGTAAACACGAACCAACATGTTAAC
 ACCTGATTGCAACTCATCTCCATTACACGTGTAAAGATCTAACATCACGAAC
 GACACCACCGCACCGTGTGGTACACGAAGAGAAGTATCACGCACCTCACGA
 GACTTGTCTCCAAAGATAGCGTGCAAGAGACGTTCTCAGCTGAAAGATCTT
 15 CTCACCCCTAGGTGTTACTTACCTACAAGAATATCACCTCTTAAACCTCAGCA
 CCAATACGGATAATCCCATTCTGCAAGGTCTTGAGGGCATCTCACCAACG
 TTTGGAATTCTCGCGAGTGATTCTCAGGTCCAA – 3'

SEQ.ID. n°13 : séquence partielle du gène *rpoB* chez *Streptococcus oralis* CIP
 20 102922^T mesurant 694 paires de bases :

5'-
 ACTCGTAAATTCCAAACGTTGGTGAAGATGCCCTAAAGACCTTGACGAAAT
 GGGTATTATCCGTATTGGTGCTGAGGTTAAAGAAGGAGATATCCTGTAGGT
 AAAGTCACACCTAACGGTGAAAAAGACCTTCTGCTGAAGAACGTCTTGCA
 25 CGCTATCTCGGAGACAAGTCTCGTGAAGTGCCTGATACTCTCTCGAGTAC
 CTCACGGTGCCGATGGTGTCTCGTGAATAGATCTGCTGGTCGTACATCGCTCA
 GGTGATGAGTTGCAATCTGGTGTGAATATGCTGGTCGTACATCGCTCA
 AAAACGTAAGATCAAGTCGGAGATAAGATGGCCGGACGTACGGAAACAAAG
 GGGTTGTCTCGTATCGTCTGTAGAAGACATGCCTACCTCCAGATGGA
 30 ACTCCAGTCGATATCATGTTGAACCCACTTGGGGTGCCTCACGTATGAATAT
 CGGTCAAGGTATGGAACCTCACCTGGTATGGCAGCCGTACTCTGGTATCC
 ACATCGCAACACCAAGTCTTGACGGAGCAAGTCGGAAGACCTTGGGACACT
 GTTAAAGAAGCAGGTATGGATAGCGATGCCAAAACAATCCTTACGATGGAC

GTACAGGTGAGCCGTTGACAACCGTGTATCAGTTGGTGTATGTACATGATC
AAACTCCA- 3'

SEQ.ID. n°14 : séquence partielle du gène *rpoB* chez *Streptococcus mutans* CIP

5 103220^T mesurant 728 paires de bases :

5' – TGTCAACCATGTGGTGAAGTTAAATCATATACATAACACCAACTGATA
CACGATTATCAAACGGCTCCCTGTGCGACCATCATAAAGAACAGTTTAGCA
TCAGAATCCATACCGGCTTACTGTTCCAAAGATCATCAGAAGTTGCT
CCGTCAAAGACAGGCGTTGCAATATGAATGCCAAATTACGAGCAGCCATACC
10 AAGATGGAGTTCCATAACTGCCAATGTCATCCGTGATGGCACCCAAAGTG
GATTAAGCATGATATCAACAGGTGTTCCATCTGGAAGATATGGCATATCTCC
ACTGGTACAATACGGGAAACGACACCCCTGTTACCATGACGTCCGGCCATCTT
ATCTCCGACCTTGATTTCACGTTTGAGCGATATAAACACGAACCAGCATGTT
AACACCTGATTGAAGTTCATCTCCATTAGCACGTGAAAGATTTCACATCACA
15 AACAAACACCGTCGCCACCATGAGGTACACGAAGAGAAGTATCAGAACCTCAC
GTGATTGTCACCAAAATGGCATGCAAGAGGGCGTCTCTGCAGAAAGATCTC
TTCTCCCTTAGGAGTCACCTTACCAACTAGAATATCACCTCTTAAACCTCAG
CACCAATGCGAATAATTCCCATTCAAGGTCTTCAGGGCATCTCACCAA
CATTTGGGATTTCACGCGTAATTCTTCAGGTCCA – 3'

20

SEQ.ID.n°15 : séquence partielle du gène *rpoB* chez *Streptococcus mitis* CIP

103335^T mesurant 730 paires de bases :

5'- TGTCAACCATGTGGGGAGTTGATCATGTAACATGACTCCGACAGA
AAACACGGTTATCAAATGGTTCACCTGTACGTCCATCGTAAAGGATTGTTTG
25 GCATCGCTATCCATACCAAGCTTCTTAAACAGTTGACCAAAGATCTCAGAACCT
GCTCCGTCAAAGACTGGTGTGCGATGTGAATACCAAGAGTACGAGCTGCCA
TCCCAAGGTGGAGTTCCATAACCTGACCGATATTACATACGTGATGGCACCCCA
AGTGGGTTAACATGATATCGACTGGAGTTCCATCTGGAAGGTAAGGCATAT
CTTCTACAGGAACGATACGAGAGACAACCCCTTATTCCGTGACGTCCGGCC
30 ATCCTATCTCCGACCTGATCTACGTTTGAGCGATGTAGACGCGAACAG
CATGTTGACACCTGATTGCAATTCACTCCATTGACGTGAAAGATCTTAAC
ATCACGAACCACACCACAGCTCCGTGGTACACGAAGAGAACAGTCACGTA
CTTCACGAGATTATCTCCGAAGATAGCGTGCAAGAGCCGTTCTCAGCTGAA
AGGTCTTCTCACCCCTAGGTGTTACTTACCTACAAGGATATCCCCTCTTAA

ACCTCAGCACCGATAACGGATAATACCCATTCTGCAAGATCTTAAGGGCATC
TTCCCCAACGTTGGGATTCTACGAGTAATTCTCAGGTCCA – 3'

SEQ.ID. n°16 : séquence partielle du gène *rpoB* chez *Streptococcus equinus*

5 CIP 102504^T mesurant 697 paires de bases :

5'-

CACTCGCGAAATTCAAACGTTGGTGAAGAAGCTCTAAAGACCTTGACGAAA
TGGGTATTATCCGTATCGGTGCTGAAGTAAAGAAGGTGACATCCTGTAGG
TAAAGTAACACCTAAAGGTAAAAAGACCTTCTGCTGAAGAGCGCCTCTC
10 ACGCAATCTCGGTGATAAAATCACGTGAAGTTCGTGATACATCACCTCGTGTA
CCACACGGTGGAGATGGTGTGTTGACGTTAAACATGCTCGTGTGTTATATCGCAC
CGGTGATGAATTACAATCAGGTGTTAACATGCTCGTGTGTTATATCGCAC
AAAAACGTAAAATCAAAGTCGGAGATAAAATGCCGGTCGTACGGTAACAA
AGGGGTTGTTCTCGTGTGTTCCAGTTGAAGACATGCCATTATCTCCAGACG
15 GAACTCCAGTCGATATCATGTTGAACCCACTGGGGTGCCTCTCGTATGAAC
ATCGGACAAGTTATGGAGCTTCACCTGGTATGGCTGCTCGTAACCTGGTAT
TCACATTGCAACACCAAGTCCTTGATGGGCAACTCTGAAGACCTTGGGATA
CAGTTAACGAAGCTGGTATGGCTAGCGACGCTAACAGACAGTTCTTACGATGG
ACGTACTGGTGAACCATTGATAACCGTGTGTCAGTTGGTGTATGTACATGA
20 TTAAACTTCAC – 3'

SEQ.ID. n°17 : séquence partielle du gène *rpoB* chez *Streptococcus constellatus*

CIP 103247^T mesurant 731 paires de bases :

5'- AGTTGTCAACCATGTGTGCAATTAAATCATATACTGACACCGACAGA
25 TACACGGTTGTCAAACGGCTGCCCGTACGACCATCATAAAGAATCGTCTGG
CATCGCTATCCATGCCTGCTCACGAACAGTATCCAAAGGTATCTGAGCTT
GCTCCGTCAAATACTGGCGTTGCTATGTGGATACCAAGGTTGCGAGCAGCCA
TACCAAGGTGAAGCTCCATAACCTGTCCGATATTACGTGATGGCACCCCA
AGTGGGTTCAACATGATGTCTACTGGTGTCCGTCGGAAAGATAAGGCATAT
30 CCTCAACTGGAACGATAACGGAAACAACCCCTTATTCCGTGGCGTCCGGCC
ATCTTATCCCCAACGCGGATCTTCGTTTGAGCAATGTAACACGCACCAAC
ATGTTGACACCAGATTGCAATTACATACCGTTCGCACGAGTAAAGATTTCAC
ATCACGGACAACCCAGCACCACCATGTGGTACACGAAGAGATGTGTCACGTA
CTTCACGAGATTATCACCGAAAATTGCATGAAGCAGGCGTCTCAGCGGAT

AAGTCTTTTCACCTTCGGCGTTACTTTACCGACAAGAATGTCGCCCTTTCA
 ACCTCAGCACCAATGCGGATAATTCCCATTCTGCAAGGTCTCTAGCGCATCT
 TCCCCAACGTTGGAATTGCGCGTAATTCTTCAGGTCCAA – 3'

5 SEQ.ID. n°18 : séquence partielle du gène *rpoB* chez *Streptococcus anginosus* CIP 102921^T mesurant 697 paires de bases :

5' –

CACGCGCGAAATTCCAAACGTCGGTGAAGATGCTTGGAGAGACCTTGACGAA
 ACGGGAATTATCCGCATTGGTGCTGAGGTAAAAGAAGGCGACATTCTGTGCG
 10 GTAAAGTAACACCGAAAGGTGAAAAAGACTTATCTGCTGAAGAACGCCTGCT
 TCATGCAATTTCGGTGATAAAATCTCGTGAAGTACGTGATACTTCCCTTCGTGT
 ACCACATGGTGGTGCAGGGGTTGTCCGTGATGTGAAAATCTTACTCGTGCG
 AACGGTGATGAATTGCAATCTGGTGTCAACATGTTGGTACGTGTTACATCGC
 TCAAAAACGGAAAATCCGTGTTGGGGATAAGATGGCTGGACGTACGGAAAC
 15 AAAGGGGTTGTTCCCGCATTGTTCCAGTTGAGGATATGCCGTATCTCCAGA
 TGGAACACCAGTTGATATTATGTTGAACCCACTGGGGTGCATCTCGTATGA
 ATATTGGTCAAGTTATGGAGCTTCACCTCGGTATGGCTCGCAACCTTGGC
 ATTACACATTGCAACACCACTATTGACGGGGCTAGCTCAGATGATCTTGGGAA
 AACCGTTCGTGAAGCTGGCATGGATAGCGATGCTAAGACAATCCATTATGAT
 20 GGCGTACTGGTGAGCCATTGATAATCGTGTATCCGTTGGTGTACATGTACAT
 GATCAAACCTCCAC – 3'

SEQ.ID. n°19 : séquence partielle du gène *rpoB* chez *Streptococcus dysgalactiae* CIP 102914^T mesurant 728 paires de bases :

25 5' – TGTCAACCATGTGGTGGAGTTAACATGTACATGACACCAACGGAT
 ACACGGTTGTCAAATGGTTCGCCAGTACGTCCATCATAAAGGACCGTCTTAGC
 ATCGCTATCCATACCAGCTCACGAACAGTGTCCAAAGGTCTCTGATGAAG
 CCCC GTCAAAGACAGGTGTTGCAATGTGAATACCAAGATTACGAGCAGCCATA
 CCAAGGTGAAGTTCCATAACCTGACCAATGTTACCGTGATGGCACCCAAAG
 30 AGGGTTCAACATGATGTCAACTGGTGTCCATCTGGAAAGGTATGGCATGTCTT
 CAACTGGTACAATACGTGAAACGACACCCCTGTTCCGTGACGACCAGCCATT
 TTATCTCCGACTTGATCTACGTTTTGAGCAATGTAAACACGCACAAGCATA
 TTAACACCTGATTGCAATTGACCGTTAGCGCGTGTAAAGATTTCACATCA
 CGAACGATACCATCACCAACCGTGAGGTACACGAAGGGACGTATACGAACCTTC

ACGTGATTATCTCCAAAGATGGCATGCAAGAGACGCTTCAGCAGAAAGGT
 CTTTTCACCTTAGGTGTGACTTACCTACTAAGATGTCGCCCTTAAACCTC
 AGCACCGATAACGGATAATTCCCATTCGTCAAGGTCTTGAGCGCTTCACC
 AACGTTGGAATTTCGCGGGTGATTCTTCAGGTCAA – 3'

5

SEQ.ID. n°20 : séquence partielle du gène *rpoB* chez *Streptococcus bovis* CIP 102302^T mesurant 728 paires de bases :

5' – TGTCAACCATTGGTGAAGTTGATCATGTACATGATAACACAGAG
 ACACGATTATCAAATGGTACCTGTACGACCGTCATAAAGAACTGTCITAGC
 10 GTCGCTATCCATACCAAGCTCACGAACAGTATCCAAAGGTCTGAAGTTG
 CCCCCGTCAAAGACTGGAGTTGCAATGTGAATACCGAGGTTACGAGCTGCCAT
 ACCAAGGTGAAGTTCCATAACTTGTCCGATATTACGAGATGGCACCCCAA
 GAGGGTTCAACATGATATCAACTGGAGTTCCGTCGGAAGATATGGCATGTC
 TTCAACAGGAACGATAACGAGAAACAACCCCTTGTTCCGTGACGACCGGCCA
 15 TTTATCTCCGACTTGTATTACGTTTGCAATGTAAACACGAACGAGCA
 TGTGACACCTGATTGCAATTACGACCGTTAGCACGTGTGAAGATTAAACA
 TCACGAACAAACACCGTCTCCACCGTGTGGCACACGAAGTGATGTATCACGTAC
 TTCACGAGATTATCACCGAAGATTGCGTGAAGAAGGCGTCTCAGCAGAAA
 GGTCTTTTACCTTAGGTGTTACCTACCTACAAGGATATCACCTCTTAA
 20 CTTCAGCACCGATAACGGATAATACCCATTCTGTCAAGGTCTTAAGAGCTTC
 CACCAACGTTGGAATTTCGCGAGTGATTCTTCAGGTCAA – 3'

SEQ.ID. n°21 : séquence partielle du gène *rpoB* chez *Streptococcus acidominimus* CIP 82.4^T mesurant 728 paires de bases :

25 5'- TTGTCAACCATTGGTGGAGCTTAATCATGTACATGACACCAACAG
 ACACACGGTTATCAAATGGTACCCAGTACGACCATCATAAAGAAATCGTTTA
 GCATCGCTGCCATTCTGCCTTTAACAGTTGACCAAGAGATCCCTTGAGCTC
 GCACCATCGAAAACCGGTGTTGCGATATGGATAACCAAGTTACGAGCAGCCAT
 ACCCAAGTGCAGTCCATAACCTGACCAATATTACGAGATGGCACCCCAA
 30 GTGGGTTCAACATGATGTCAACTGGTGTCCATCTGGAAGATATGGCATGTCT
 TCAACTGGTACAATACGAGAAACGACACCCCTGTTACCGTGACGACCGGCCAT
 CTTATCTCCGACCTTAATCTGCGTTTGAGCGATATAACACACGTACCAAGCAT
 ATTAACACCAAGACTGTAGCTCATCACCATTAGCACCGTAAAGATTTCACATC
 ACGAACAAACACCATTGCACCGTGTGGCACACGTAGAGAGGTATACGTACTT

CACGTGATTGTACCGAAGATAGCATGCAAGAGACGCTCCTCAGCAGAAAG
 ATCTTTCACCTTGTTGGTGTACCTTACCAACAAGAATATCGCCTCTTAACCT
 TCTGCACCGATAACGGATAATACCCATTCTGTCAAGGTCTTGAGGGCTCTC
 ACCAACGTTGGAATTTCACGAGTAATTCTTCAGGTCA – 3'

5

SEQ.ID. n°22 : séquence partielle du gène *rpoB* chez *Streptococcus agalactiae* CIP 103227^T mesurant 733 paires de bases :

5' – TGAGTTGTACCAACCATGTGGTGAAGTTGATCATGTACATGACACCAA
 CTGACACACCGTTATCGAATGGTTACCAAGTACGACCATCATAAAGAACAGTC
 10 TAGCATCTGAATCCATACTGCTTCTGAACAGTTCCAAAGGTCTCTGAA
 GAAGCCCCATCAAAGACTGGCGTTGCAATATGAATACTAAATTACGAGCAGC
 CATACTAAATGAAGCTCCATAACTTGTCCGATATTACGTGATGGCACCCC
 AAGTGGGTTCAACATGATATCAAATGGCGTTCCATCTGGTAAGTAAGGCATAT
 CTTCAACAGGAACAATACTGAGACGACACCTTGTCCGTGACGACCGGCC
 15 ATCTTATCACCGACTTGATTACGTTTACGTTTGAGCGATATAACGCGGACAAG
 CATATTAACACCTGATTGCAATTACCATTTGCACGAGTAAAGATTTAAC
 'GTCACGAACTACTCCATGCCACCGTGAGGTACACGTAGTGAAGTATCAGAA
 CTTCACGTGATTATCACCAAAATGGCATGCAAGAGACGTTCTCAGCAGAT
 AAGTCCTTTCACCCCTAGGTGTTACCTTACCAACAAGAATGTCACCTCTTT
 20 ACCTCAGCACCAATGCGGATAATTCCCATTACGTGAGATCACGTAGTGAATC
 TTCACCAACATTGGATTTCACGAGTAATTCTTCAGGTCCA – 3'

SEQ.ID. n°23 : séquence partielle du gène *rpoB* chez *Streptococcus difficile* CIP 103768^T mesurant 714 paires de bases :

25 5'- TTGTACCAACCATGTGGTGAAGTTGATCATGTACATGACACCAAATGAC
 ACACGGTTATCGAATGGTTACCAAGTATGACCATCATAAAGAACAGTCTTAGCAT
 CTGAATCCATACTGCTTCTGAACAGTTCCAAAGGTCTCTGAAGAACGGCCC
 ATCAAAGACTGGCGTTGCAATATGAATACTAAATTACGAGCAGCCATACCTAAA
 TGAAGCTCCATAACTTGTCCGATATTACGTGATGGCACCCCAAGTGGGTTCA
 30 ACATGATATCAAATGGCGTTCCATCTGGTAAATAAGGCATATCTCAACAGGAAC
 AATACGTGAGACGACACCTTGTCCGTGACGACCGGCCATCTTATCACCGACT
 TTGATTTACGTGTTGAGCGATATAACGCGGACAAGCATATTAAACACCTGATT
 GCAATTACATCACCAATTGCACGAGTAAAGATTAAACGTACGAACACTCCATC
 GCCACCGTGAGGTACACGTAGTGAAGTATCACGAACCTCACGTGATTATCACCA

AAAATGGCATGCAAGAGACGTTCTCAGCAGATAAGTCCTTTCACCCCTAGGCG
 TTACCTTACCAACAAGAATGTCACCTTCTTTACCTCAGCACCAATGCGGATAATT
 CCCATTTCATCGAGATCACGTAGTGAATCTCACCAACATTTGGAATTTCACGAG
 TA - 3'

5

SEQ.ID. n°24 : séquence partielle du gène *rpoB* chez *Streptococcus intermedius*

CIP 103248^T mesurant 728 paires de bases :

5'- TGTATCAACCAGTGGTGAAGCTTAATCATGTACATGACACCAACGGAC
 ACACGGTTATCAAACGGTTCGCCAGTACGTCCATCATAAAGGATTGTCTTAGC
 10 ATCGCTATCCATACCTGCTCACGAACGGTTCCAAAGATCATCTGAGCTAGC
 TCCGTCAAAGACTGGCGTTGCAATGTGGATACCAAGTTGCGAGCAGCCATAC
 CGAGGTGCAATTCCATAACTGTCCGATATTACATACGTGACGGCACCCAAAGA
 GGATTCAACATGATATCAACTGGTGTCCGTCTGGAAGATAACGGCATATCCTC
 AACTGGAACAATGCGGGAAACAACCCCTTGTTCGTGGCGTCCGGCCATCT
 15 TATCTCCAACGCGGATTTCGTTTGAGCGATATAAACACGTACCAACATGT
 TGACACCGGATTGCAATTCAATCACCCTCGCACGAGTAAAGATTTCATCAC
 GGACAACACCTGCACCACCGTGTGGTACACGAAGGGAGGTATACGCACCTC
 ACGAGACTTATCACCAAAATGCATGAAGCAAGCGTTCTCAGCGGATAAAAT
 CTTTCACCTTCGGCGTTACCTACCGACAAGAATGTCGCCCTTTACCTC
 20 AGCACCAATGCGGATAATTCCATCTCGTCAAGGTCTCTCAAAGCATCTCCCC
 GACGTTGGAATTGCGCGTGATTCTTCAGGTCCA - 3'

SEQ.ID. n°25 : séquence partielle du gène *rpoB* chez *Streptococcus equi* CIP

102910^T mesurant 728 paires de bases

25 5'- TGTATCAACCAGTGGTGAAGCTTAATCATATACATGACACCAACTGAC
 ACACGATTATCAAACGGCTCACCAAGTACGGCCATCATAAAGAACAGTCTTAGC
 ATCGCTATCCATACCTGCTCACGAACAGTTCCAAAGGTCTCAGACGTAGC
 TCCGTCAAAGACCGGTGTGCGATATGGATACCCAAATTACGAGCAGCCATAC
 CTAGGTGAAGCTCCATAACCTGTCCAATGTTACATACGAGACGGCACCCAAAGA
 30 GGGTCAGCATGATGTCAACAGGGTTCCGTCTGGCAGATATGGCATATCCT
 CAACCGGTACAATACGTGAGACGACACCCCTGTTACCATGACGCCGGCCATT
 TTATCTCCGACCTTGATTTCAGCTTTGAGCAATGTAAACACGCACCAAGCATA
 TTAACACCTGATTGAAGCTCATCACCATTGCGCGTGTAAAGATCTCACATCA
 CGTACAATCCCGTACCAACCAGTGGAACACGTAACGAGGTATACGAACCTC

ACGTGATTATCACCAAAGATAGCATGCAGGAGACGTTCTCAGCAGAAAGG
TCTTTTCACCCCTAGGAGTTACCTTACCAACAAGAATATGCCCTCCTGACC
TCTGCACCGATAACGGATAATACCCATTTCATCAAGGTCCCTGAGGGCTTCTCA
CCAACGTTGGCACTTCACGTGTGATTCTCAGGTCCA – 3'

5

SEQ.ID. n°26 : séquence partielle du gène *rpoB* chez *Enterococcus gallinarum*
CIP 103013^T mesurant 694 paires de bases :

5'-

CACTCGTAAATCCGAATGTCGGGAAAGACGCATTGAAAGATCTAGACGAA
10 ATGGGTATCATCCGCATTGGTGCAGAAGTCAAAGATGGCGATCTGTTGGTG
GTAAAGTAACGCCTAAAGGGTAACGGAACATCTGCAGAAGAACGCTTGCT
TCATGCAATCTTGGTAAAAAGCCCGCGAAGTCCCGCGATACTCTCTGCGCG
TACCTCACGGTGGTGGCGGAATCGTCATGATGTGAAAATCTTACCCGCGAA
GCTGGCGATGAATTGTCAACCAGGTGTCAATATGCTCGTGTGTATATCGT
15 TCAAAAACGGAAAATCCATGAAGGGGATAAAATGGCCGGCGTCACGGAAAT
AAAGGGTCGTTCTCGCATTATGCCAGAAGAACATGCCCTTACCGAGA
CGGTACACCAGTTGATATCATGTTGAACCCATTAGGGTGCCTCACGGATGA
ACATTGGACAAGTATTGGAATTACACTTAGGAATGGCTGCCCGCCAATTAGGA
ATCCACGTGGCTACACCAGTCTTGATGGTGCCAGCGATGAAGATGTCTGGG
20 CAACAGTGCAGAAGCCGGCATGGCTAGCGACGCCAAACCGTTTGATGA
TGGCCGTACTGGAGAACCATTTGATGGTCGAATCTCCGTAGGTGTATGTATA
TGATCAAATTGGCC – 3'

SEQ.ID. n°27 : séquence partielle du gène *rpoB* chez *Enterococcus*
25 *casseliiflavus* CIP 103018^T mesurant 727 paires de bases :

5'- TGTCAACCATGTGGGCCATTGATCATGTACATGACACCAACGGAG
ATGCGGCCATCAAATGGTCCGGTACGTCCGTCAAAGCACTGTTGGC
ATCGCTGGCCATTCCCTGCTTCAGCAACCGTTGCCAAACATCTCATCGCTGGC
TCCATCAAAGACTGGTGTGCCACGTGAATGCCTAATTGACGCGCAGCCATT
30 CTAAGTGTAACTCTAATACTGTCCAATGTTCATCCGAGAACGGTACCCCTAATG
GGTTCAGCATGATATCGACTGGTGTGCCATCTGGTAAGAAAGGCATGTCTTCT
TCTGGCATAATGCGAGAACGACCCCTTGTGTTCCGTGACGTCCGGCCATT
ATCCCCCTCATGGATTTCGTTTGAACGATATAACGCGAACCAAGCATGTT
CACACCTGGTGACAATTCATGCCAGCTCGCGGGTAAAGATTGACATCGT

GGACGATTCCGCCGCCGTGAGGCACCGTAGAGAAGTGTACGCACITC
 5' GCGGGCTTTCACCAAAGATTGCGTGCAACAAACGCTCTGCTGAAAGTT
 CCGTTACCCCTTGGCGTGACTTCCAAACAAGCAGATGCCATCTGACTT
 CCGCACCAATGCGGATAATGCCATTGTCTAGGTCTTCAACGCGTCTCCC
 5 AACGTTGGGATTTCGCGAGTGATTCTCAGGTCCA – 3'

SEQ.ID. n°28 : séquence partielle du gène *rpoB* chez *Enterococcus saccharolyticus* CIP 103246^T mesurant 721 paires de bases :

5'- TGTCAACCATGTGGCAAGTTAACATGTACATTACCCAACAGAG
 10 ATACGACCATCGAATGGTCACCGTACGTCCGTACATAAACAGAGTTTCGC
 ATCGCGGCCATGCCGCTCGCGAACTGTTCCATACGTACATCTGATGC
 ACCATCAAATACTGGTAGCTACATGGATGCCTAACTGACGTGCAGCCATCC
 CTAAGTGTAAATTCCAATACTGTCCGATGTTACAGAGATGGTACTCCTAGT
 GGGTTAACATGATATCAAACGGTGCCGTCTGGTAAGAACATGGCATGTCITC
 15 TTCTGGCATAATGCGAGAGACAACCCCTTGTACCATGACGTCCGCCATTTC
 ATCTCCTCGTGAATCTACGTTGCACGATAAACACGAACAAACATGTT
 CACACCTGGAGATAATTGTCGCCTGCTCACGGTAAAGATTTAACATCGT
 GAACGATAACGCCACCGCCGTGAGGAACACGTAATGATGTACGTACTTCA
 CGTGTCTTCACCGAAGATTGCGTGAATAGACGTTCTGCAGATAATT
 20 GGTACCCCTTACGGAGTGACTTACCTACTAATAAGTCGCCATCTGTACTTC
 GGCACCGATAACGGATAATACCCATTGTCTAAGTCTTAAATGCGTCTCCC
 AACGTTAGGAATTTCGCGTGTATTCTCAG – 3'

SEQ.ID. n°29 : séquence partielle du gène *rpoB* chez *Enterococcus faecium* CIP 103014^T mesurant 727 paires de bases :

5'- TGTCAACCATGTGAGCAAGTTGATCATGTACATCACACCGACAGAC
 ACACGTCCATCAAATGGTCACCTGTACGTCCGTACAGAACAGTTTCGC
 ATCGCTGGCCATACCGGCTCACGAACCTGTTCCATACGTCTCATCACTTGC
 ACCATCAAATACTGGCGTGCTACGTGGATACCTAACTGACGTGCAGCCATAC
 30 CCAAGTGTAAATTCCAATACTGCCGATGTTACGTGAAGGCACCCCTAAA
 GGATTCAAGCATGATATCGATTGGTGTCCATCAGGTAGGAATGGCATATCTC
 TTCCGGCATAATACGGGATAACAACCCCTTATTCCGTGACGACCGGCCATTTC
 ATCCCCCTCATGGATTACGTTTGAACGATATAAACACGAACAAACATGTT
 TACGCCTGGTGACAATTCATCTCCAGCTCACGAGTAAAGATTTCACATCGT

GAACGATACGCCGCCATGTGGTACACGTAATGATGTATGCCGGACTCA
 CGAGCTTTTCGCCAAAGATCGCATGCAATAGACGTCTCTGCAGATAATTCT
 GTTACCCCTTTGGCGTGACTTCCCTACAAGCAAATGCCATCTGGACTTCT
 GCACCAATACGGATGATACCCATTCTGTCTAAATCTTTAATGCGTCTCCGA
 5 CATTAGGGATTTCGCGTGTGATTCTTCAGGTCCA – 3'

SEQ.ID. n°30 : séquence partielle du gène *rpoB* chez *Enterococcus faecalis* CIP 103015^T mesurant 724 paires de bases :

5'- TGTCATCAACCATTGTGGGCTAATTAAATCATATACATGACACCAACGGAA
 10 ATACGGTTATCAAATGGTCACCTGTACGTCCATCGTAAAGAACTGTTTAC
 ATCGCTAGCCATACCAAGCTCACGAACAGTTCCAAACGTCTCATCGGTTGC
 CCCATCGAAAACAGGTGTTGCGACGTGAATACCTAATTGGCGAGCAGCCATAC
 CTAAGTGTAAATTCAAGTACTTGTCCGATATTACATACGAGAAGGTACCCCTAAT
 GGGTTCAACATGATATCAACAGGTGTTCCGTCAAGGTAAAGAATGGCATATCTC
 15 TTCCGGCATAATACGGGAAACAACCCCTTATTTCGTGACGTCCGCCATTTC
 ATCTCCTTCGTGAATTTCACGTTTGAACGATATAGACACGAACAACTATGTT
 GACACCTGGTATAATTACATGCCAGCTCACGAGTAAAGATTTCACATCAT
 GAACGATACGCCGCCACCGTGAGGTACACGGAGAGACGTATACGAACCTC
 GCGGGCTTTTCCCGAAGATTGCGTGTAAATAACGTCTGCAGATAATT
 20 CTGTGACCCCTTAGGTGTGACTTCCCAACTAGTAAGTCGCCATCTGAACCT
 CAGCACCAATGCGGATAATCCCCATTCTGTCTAAGTCCTAACGCGTCTCC
 AACGTTTGGAATTTCACGGTATTCTTCAGGTCA – 3'

SEQ.ID. n°31 : séquence partielle du gène *rpoB* chez *Enterococcus avium* CIP 103019^T mesurant 570 paires de bases :

5'- GTCCATCATAAAGAACGGCTTACGCATCTGCTGCCATACGAGCTCACGA
 ACTGTTCCAAACATCGCTATCTGCGCACCATCGAAGACTGGTGTGCAAC
 ATGGATACCTAGTGGCGAGCCGCCATTCCAAGTGTAAATTCAAACACTTGT
 CGATGTTCATCCGAGATGGCACACCTAATGGGTCAACATGATATCAACTGGC
 30 GTACCGTCTGGTAAGAAAGGCATGTCTCTGGCATAATGCGAGAAACGA
 CCCCTTATTCCGTGACGGCCGGCATTTATCCCCTCATGAATCTACGTT
 TTTGCACGATGTACACGCGCACTAACATATTACACCTGGAGATAATTAC
 CTGCTTCACGAGTAAAGATCTCACATCGTGAACGATCCGCCACCAGTC
 GGTACACGAAGAGATGTATACGAACCTCACGAGCCTTCACCAAAGATCGC

ATGCAACAAACGTTCTCAGCTGATAATTCTGTTACCCCTTACGGAGTGACTTT
ACCAACTAATAAATCACCATCATGAACITCAGCACCAATAAC -3'

SEQ.ID. n°32 : séquence partielle du gène *rpoB* chez *Abiotrophia defectiva* CIP

5 103242^T mesurant 732 paires de bases :

5'- GAAGTTGTCATCAACCATGTGGGCCAACCTAATCATGTACATAACCCCAA
CAGAGACTTTACGGTCAAATGGTTACCGGTTCGACCATCATATAAGATAGTC
TTAGCGTCAGCTCTAAGCCGGCTCCCTAACTGTTCCAGACATCTTCTCA
CTAGCACCGTCAAAGACAGGTGTTGCAATCTGATGCCATTGCGAGCAGC
10 CATCCCCAAGTGTAACTCTAGGACTTGCCGATGTTACATGGGATGGAACCC
CTAATGGGTCAACATGATATCAACTGGGTACCATCTGGTAAGAATGGCATA
TCTTCTCCGGCATGATAAGGGAGACAACCCCTTGTACCGTGACGACCGGC
CATCTTATCCCCTTCATTGATTACGTTTGTACGATGTAGACGCGGACTAG
CTTGTGACACCTGGTGCCAATTCTCGCCAGCTCGCGGGTAAAGATTTAA
15 CGTCGTGGACAATCCCGCCCCCGCCGTGTGGCACACGCAAGGAAGTATCACG
TACTTCACGCGCCTCTCACCGAAGATAGCATGGAGCAAGCGTTCTCCGCAG
ACAACCTGGTCACACCTTGGTGTACCTTACCAACTAAGATATGCCGTCTT
TTACTTCCGCCCGATAACAGATAATCCGCTTGGTCTAAGTACTTGAGGGCA
TCTTCGGACACGTTGGAATTCTCGGTGTAATTCTTCAGGTCA - 3'

20

SEQ.ID. n°33 : séquence partielle du gène *rpoB* chez *Gemella morbillorum* CIP

81.10^T mesurant 727 paires de bases :

5'- TGTCAACCATGTGTGCAAGTTATCATGTACATTACCCCTACAGATAC
ACGGCTATCAAATGGCTCACCTGTACGTCCGTATAAAGAACTGTCTAGCAT
25 CTTAGCCATTCCAGCTTCCGCAACTGTAGACCAAAACATCTCATCAGTAGCAC
CATCGAATACTGGTAGCTACGTGGATTCCAAGTTGTACAGCCATACCT
AAAGTGTAGCTTAATACTTGTCCAATGTTACAGAGATGGAACCCCAAGTGG
GTTAACATTACGTCAACTGGTAGCATCTGGTAGGTAAGGCATATCTTCTT
CTGGTAAGATATTGAGATAACCCCTTGTACCGTGACGACCGGCCATTAA
30 TCTCCTACACGAATTACGTTTGGACGATAAATACACGAACAAAGTTCATTT
ACACCGTTAGGTAATTCTACGACCCATCTCACGTTAAAGATTAAACATCAGCA
ACTACTCCATCAGCACCGTGAGGTACACGTAATGAAGTATCACGTACTCTTAA
GATTAGCTCAAAGATAGCATATAATAATTCTCTGGAGTTGTACGTT
AATCCATTCTGGTGTAACTTACCTACTAAAATATCTCCATCTTAACTTCAGCC

CCAATACGAATGATTCCCTCGTGCATCTAAGTTCTAAGTGCATTTCACCCCTAC
GTITGGAATCTCACGAGTAATTCTTCAGGTCA – 3'

SEQ.ID. n°34 : séquence partielle du gène *rpoB* chez *Gemella haemolysans* CIP

5 101126^T mesurant 726 paires de bases :

5'- TGTCAACCAGTGTGCAAGTTAACATGTACATTACCCCTACAGATA
CACGGCTATCAAATGGCTCACCTGTACGTCCGTATAAAGAACTGTCTTAGCA
TCTTAGCCATTCCAGCTTCCGCAACTGTAGACCAAACATCTCATCAGTAGCA
CCATCGAATACTGGTAGCTACGTGGATTCCAAGTGTAGCAGCCATACC
10 TAAGTGTAGCTCTAACACTGTCCAATGTTACAGAGATGGAACCCCAAGTG
GGTTAACATTACGTCAACTGGTAGCATCTGGTAGGTAAGGCATATCTCT
TCTGGTAAGATATTGAGATAACCCCTTGTACCGTGACGACCGGCCATT
ATCTCCTACACGAATTTACGTTGGACGATAAAATACACGAACAAAGITCATT
TACACCGTTAGGTAAATTCAAGCACCACGTTACGTTAAAGATTAAACATCAGC
15 AACTACTCCATCAGCACCGTGAGGTACACGTAATGAAGTATCACGTACTCTT
AGATTAGCTCCAAAGATAGCATATAATAATTCTCTGGAGTTGTTAGT
TAATCCCTTCGGTGTAACTTACCTACTAAAATATCTCCATCTTAACCTCAGC
CCCAATACGAATGATTCCCTCGTGCATCTAAGTTCTAAGTGCATTTCACCC
GTITGGAATCTCACGAGTATTCTTCAGGTCCA – 3'

20

SEQ.ID. n°35 : séquence partielle du gène *rpoB* chez *Granulicatella adjacens*
CIP 103243^T mesurant 719 paires de bases :

5'- CATCAACCAGTGTGAGCAAGTTGATCATGTACATAACCCCTACTGACACA
CGGTTATCGAATGGTCCCCGTACGTCCATCATAGAATTGTTTCGCATCA
25 CGAGCCATACCCGCTCTGCAACAGTTCCCCATACGTCTCATCTGCGCACCA
TCGAATACTGGTGTGCGATGTAAATACCTAACGAGCAGCCATCCCTAA
GTGTAACCTAACACTGTCCGATGTTACACGTGAAGGTACCCCTAATGGGT
TTAACATGATGTCAACTGGTGTCCATCTGGTAAGAATGGCATATCTCTCC
GGCATAATACGGGAAACAACCCCTTATTACCGTGACGTCCGGCCATCTTATC
30 CCCTTCATTGATTTACGTTGTACAATATACACGAACTAATTGTTACG
CCAGGTGCTAATTCACTCACCTGCTGCACGTGTGAATAACACGTACATCACGGAC
AATACCGCCACCGCCGTGAGGTACACGTAGAGATGTGTACGAACCTCACGA
GCTTTTCACCGAAGATTGCGTGTAAATAAACGTTCCCTGGTGTATTGTTCTGTT
AACCCCTTAGGAGTTACTTACCAACTAAGATGTCAACCACCTTAACCTCGGCA

CCGATACGAATAATTCCGTCTGCGTCTAGGTTCTCAATGCGTCTCCAACGT
TTGGAATCTCACGAGTAATTCTCAGG – 3'

Dans les séquences ci-dessus, le nucléotide M désigne A ou C, le nucléotide R désigne A ou G, le nucléotide W désigne A ou T, le nucléotide Y désigne C ou T et le nucléotide N désigne A, T, C ou G.

Dans les séquences ci-dessus, les références CIP se rapportent à des dépôts à la Collection Nationale de Culture des Microorganismes (CNCM) de l'Institut Pasteur à Paris (France).

10 Exemple 3. Identification en aveugle d'une collection de 20 souches bactériennes comprenant 10 souches de bactéries appartenant aux genres *Streptococcus* et genres apparentés.

Une collection de vingt souches appartenant aux espèces bactériennes suivantes: *Streptococcus pyogenes*, *Streptococcus sanguis*, *Granulicatella adjacens*, *Abiotrophia defectiva*, *Enterococcus avium*, *Enterococcus faecalis*, *Gemella haemolysans*, *Gemella morbillorum*, *Streptococcus equi*, *Streptococcus anginosus*, *Staphylococcus aureus*, *Pseudomonas oleovorans*, *Mycobacterium avium*, *Bacillus cereus*, *Acinetobacter anitratius*, *Corynebacterium amycolatum*, *Klebsiella terrigena*, *Pasteurella*, *Lactobacillus rhamnosus*, *Staphylococcus*, a été codée de façon à réaliser une identification moléculaire en aveugle (l'expérimentateur ne connaissant pas a priori l'identité des souches) des souches selon le procédé décrit dans la présente demande de brevet. L'extraction des acides nucléiques ainsi que l'amplification du fragment du gène *rpoB* ont été réalisées comme décrites dans l'exemple n°2 en incorporant des amorces consistant dans des mélanges de 4 oligonucléotides qui ont des séquences consistant dans les séquences SEQ ID N°6 (comme amorce 5') et SEQ ID N°7 (comme amorce 3') avec N représentant l'inosine, dans une amplification PCR (Fig. 1). Le séquençage de ces 10 amplifiats a été réalisé en incorporant dans la réaction de séquençage les amorces SEQ ID N°6 et SEQ ID N°7 comme décrit dans l'exemple n°2 et la comparaison des séquences obtenues avec les séquences SEQ.ID n° 1 à 5 et 8 à 35 a permis d'identifier les dix souches amplifiées comme étant *Streptococcus pyogenes*, *Streptococcus sanguis*, *Granulicatella adjacens*, *Abiotrophia defectiva*, *Enterococcus avium*,

Enterococcus faecalis, *Gemella haemolysans*, *Gemella morbillorum*, *Streptococcus equi*, *Streptococcus anginosus*. Le décodage de ces 10 souches a montré 100% de concordance entre l'identification moléculaire selon le procédé faisant l'objet de la présente invention et l'identification établie 5 antérieurement par les méthodes phénotypiques standard. Ce résultat illustre la spécificité du jeu d'amorces SEQ ID N°6/SEQ ID N° 7 ~~utilisé pour ce travail~~.

Les autres bactéries choisies pour ce qu'elles sont fréquemment isolées dans les prélèvements cliniques humains ou animaux susceptibles de contenir également des bactéries du genre *Streptococcus*, n'ont pas été amplifiées, 10 démontrant ainsi la spécificité des amorces utilisées pour le genre *Streptococcus* et dits 4 genres apparentés dans les conditions d'utilisation pour la détection des bactéries du genre *Streptococcus* et dits 4 genres apparentés selon l'invention par rapport aux bactéries d'un autre genre.

Sur la figure 1 sont représentés les produits d'amplification PCR obtenus 15 à partir de dix souches bactériennes codées, comportant 7 souches appartenant au genre *Streptococcus* et lesdits 4 genres apparentés (colonnes 2, 3, 4, 7 -11) et 3 souches bactériennes de genres bactériens autres que *Streptococcus* et lesdits 4 genres apparentés (colonnes 5,6 et 12). Les colonnes 1 et 13 représentent le marqueur de poids moléculaire. Les produits d'amplification sont 20 obtenus après incorporation des amorces SEQ ID N°6 et SEQ ID N° 7 décrits ci-dessus et sont visualisés par coloration au bromure d'éthidium après électrophorèse sur un gel d'agarose.

REVENDICATIONS

1. Gène ou fragment de gène *rpoB* d'une bactérie du genre *Streptococcus* et des 4 genres apparentés *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella*, caractérisé en ce qu'il comprend une séquence choisie parmi 5 les séquences SEQ.ID. n° 8 à 35 dans lesquelles :
 - le nucléotide K représente T ou G,
 - le nucléotide M représente A ou C,
 - le nucléotide R représente A ou G,
 - le nucléotide W représente A ou T,
 - 10 - le nucléotide Y représente C ou T,
 - le nucléotide N représente A, T, C, G ou I, et les séquences inverses et séquences complémentaires ainsi que les séquences présentant au moins 98,7% d'homologie, à l'exclusion des séquences SEQ ID n°11, 12, 14 et 22.
- 15 2. Gène *rpoB* d'une des bactéries *Streptococcus anginosus*, *Streptococcus equinus*, *Abiotrophia defectiva* et *Enterococcus faecalis* selon la revendication 1, caractérisé en ce qu'il correspond à l'une des séquences choisie parmi les séquences SEQ.ID. n° 1 à 3 et SEQ ID n°5, dans lesquelles :
 - le nucléotide K représente T ou G,
 - 20 - le nucléotide M représente A ou C,
 - le nucléotide R représente A ou G,
 - le nucléotide W représente A ou T,
 - le nucléotide Y représente C ou T,
 - le nucléotide N représente A, T, C, G ou I,et les séquences inverses et séquences complémentaires, ainsi que les séquences présentant au moins 98,7% d'homologie.
- 25 3. Fragment d'un gène *rpoB* d'une bactérie du genre *Streptococcus* et des 4 genres apparentés *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella*, caractérisé en ce que sa séquence est comprise ou consiste dans l'une des 30 séquences SEQ.ID. n° 8 à 35, dans lesquelles :
 - le nucléotide K représente T ou G,
 - le nucléotide M représente A ou C,
 - le nucléotide R représente A ou G,

- le nucléotide W représente A ou T,
- le nucléotide Y représente C ou T,
- le nucléotide N représente A, T, C, ou G
et les séquences inverses et séquences complémentaires, ainsi que les
5 séquences présentant au moins 98,7% d'homologie.

4. Oligonucléotide caractérisé en ce qu'il comprend une séquence spécifique d'une espèce d'une bactérie du genre *streptococcus* et dits genres apparentés, de préférence d'au moins 20 nucléotides consécutifs, de préférence encore au moins 30 nucléotides consécutifs inclus dans l'une des dites
10 séquences SEQ ID n°8 à 35., dans lesquels :

- le nucléotide K représente T ou G,
- le nucléotide M représente A ou C,
- le nucléotide R représente A ou G,
- le nucléotide W représente A ou T,
15 - le nucléotide Y représente C ou T,
- le nucléotide N représente A, T, C, ou G
et les séquences inverses et séquences complémentaires, ainsi que les :
séquences présentant au moins 98,7% d'homologie.

5. Utilisation d'un gène, fragment de gène ou oligonucléotide selon
20 l'une des revendications 1 à 4 à titre de sonde d'espèce d'une bactérie du genre *streptococcus* et dits genres apparentés.

6. Oligonucléotide caractérisé en ce qu'il comprend une séquence d'au moins 8, de préférence au moins 12, de préférence encore 18 à 35 motifs nucléotidiques, dont au moins une séquence de 8 motifs nucléotidiques
25 consécutifs inclus dans l'une des séquences SEQ.ID. n° 6 et 7 suivantes :

- SEQ ID N° 6 : 5' AARYTNGGMCCCTGAAGAAAT-3', et
- SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3',

dans lesquelles :
- N représente l'inosine ou l'un des 4 nucléotides A, T, C ou G,
30 - R représente A ou G,
- M représente A ou C, et
- Y représente C ou T,
et les séquences inverses et séquences complémentaires.

7. Mélange d'oligonucléotides, caractérisé en ce qu'il est constitué d'un mélange équimolaire d'oligonucléotides selon la revendication 6, ayant tous une séquence différente et comprenant tous une séquence incluse dans SEQ ID n°6 ou tous une séquence incluse dans SEQ ID n°7.

5 8. Mélange d'oligonucléotides, caractérisé en ce qu'il est constitué d'un mélange équimolaire de 32 oligonucléotides selon la revendication 7, de séquences différentes comprenant chacune au moins 15, de préférence au moins 18 motifs nucléotidiques consécutifs, inclus dans la séquence suivante :

- SEQ ID N° 6 : 5' AARYTNGGM CCTGAAGAAAT-3',

10 dans laquelle :

- R représente A ou G,
- Y représente C ou T,
- M représente A ou C, et
- N représente A, T, C ou G,

15 et les séquences inverses et séquences complémentaires.

9. Mélange d'oligonucléotides, caractérisé en ce qu'il est constitué d'un mélange équimolaire de 8 oligonucléotides selon la revendication 7 de séquences différentes comprenant chacune au moins 15, de préférence au moins 18 motifs nucléotidiques consécutifs inclus dans la séquence suivante :

20 - SEQ ID N° 6 : 5' AARYTNGGM CCTGAAGAAAT-3',

dans laquelle :

- R représente A ou G,
- Y représente C ou T,
- M représente A ou C, et
- N représente l'inosine,

25 et les séquences inverses et séquences complémentaires.

10. Mélange d'oligonucléotides, caractérisé en ce qu'il est constitué d'un mélange équimolaire de 16 oligonucléotides selon la revendication 7, de séquences différentes comprenant chacune au moins 15, de préférence au moins 21 motifs nucléotidiques consécutifs inclus dans la séquence suivante :

30 - SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3',

dans laquelle :

- R représente A ou G, et

- N représente A, T, C ou G,
et les séquences inverses et séquences complémentaires.

11. Mélange d'oligonucléotides, caractérisé en ce qu'il est constitué d'un mélange équimolaire de 4 oligonucléotides selon la revendication 7, de 5 séquences différentes comprenant chacune au moins 15, de préférence au moins 21 motifs nucléotidiques consécutifs inclus dans la séquence suivante :

- SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3',
dans laquelle :

- R représente A ou G, et
10 - N représente l'inosine,
et les séquences inverses et séquences complémentaires.

12. Mélange d'oligonucléotides selon l'une des revendications 7 à 11, caractérisé en ce que lesdites séquences consistent dans les séquences SEQ.ID. n° 6 et 7 dans lesquelles, de préférence, N représente l'inosine, et les 15 séquences inverses et séquences complémentaires.

13. Utilisation d'un oligonucléotide ou mélange d'oligonucléotides selon l'une des revendications 6 à 12, à titre d'amorce d'amplification ou sonde de genre d'une bactérie du genre *Streptococcus* et dits genres apparentés.

14. Procédé de détection par identification moléculaire d'une bactérie 20 de l'une des espèces du genre *Streptococcus* et des 4 genres apparentés *Enterococcus*, *Gemella*, *Abiotrophia* et *Granulicatella*, caractérisé en ce qu'on utilise :

- un gène ou fragment de gène *rpoB* ou un oligonucléotide selon l'une des revendications 1 à 4 ainsi qu'un gène ou fragment de gène *rpoB* d'une bactérie 25 *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* et *Streptococcus agalactiae* comprenant une séquence telle que décrite respectivement dans les séquences SEQ. ID n°11, 12, 14 et 22, ainsi que les séquences inverses et séquences complémentaires, et séquences présentant au moins 98,7% d'homologie, et/ou
30 - au moins un oligonucléotide ou mélange d'oligonucléotides selon l'une des revendications 6 à 12.

15. Procédé selon la revendication 14 dans lequel on cherche à détecter la présence d'une bactérie du genre *streptococcus* ou desdits 4 genres apparentés, caractérisé en ce qu'il comprend les étapes dans lesquelles:

1- on met en contact au moins une sonde de genre comprenant un dit mélange d'oligonucléotides selon l'une des revendications 7 à 12, avec un échantillon contenant ou susceptible de contenir des acides nucléiques d'au moins une telle bactérie du genre *Streptococcus* et ses dits 4 genres apparentés, et

2- on détermine la formation ou l'absence de formation d'un complexe d'hybridation entre ladite sonde de genre et les acides nucléiques de l'échantillon, et on détermine ainsi la présence de ladite bactérie dans l'échantillon s'il y a formation d'un complexe d'hybridation.

16. Procédé selon la revendication 14, caractérisé en ce qu'il comprend les étapes dans lesquelles :

15 1- on met en contact les amores d'amplification comprenant desdits mélanges d'oligonucléotides selon l'une des revendications 7 à 12, avec un échantillon contenant ou susceptible de contenir des acides nucléiques d'au moins une telle bactérie du genre *Streptococcus* et lesdits 4 genres apparentés, et avec :

20 - comme amorce 5', un dit mélange d'oligonucléotides comprenant une séquence incluse dans la séquence SEQ.ID. n°6, de préférence consistant dans ladite séquence SEQ ID N°6 complète, ou une dite séquence complémentaire selon l'une des revendications 7, 8, 9 ou 12, et

25 - comme amorce 3' un dit mélange d'oligonucléotides comprenant une séquence incluse dans la séquence SEQ.ID. n°7, ou de préférence consistant dans ladite séquence SEQ ID N°7 complète ou une séquence complémentaire selon l'une des revendications 7, 10, 11 ou 12

30 2- on réalise une amplification d'acides nucléiques par réaction de polymérisation enzymatique et on détermine l'apparition ou l'absence d'un produit d'amplification, et on détermine ainsi la présence de ladite bactérie dans l'échantillon si un produit d'amplification est apparu.

17. Procédé selon la revendication 14 ou 16, caractérisé en ce qu'on cherche à détecter spécifiquement une espèce donnée d'une bactérie du groupe *Streptococcus* et lesdits 4 genres apparentés choisis parmi les espèces :
Streptococcus mutans, *Streptococcus oralis*, *Streptococcus pneumoniae*,
5 *Streptococcus pyogenes*, *Streptococcus salivarius*, *Streptococcus sanguinis*,
... *Streptococcus suis*, *Streptococcus acidominimus*, *Streptococcus agalactiae*,
Streptococcus anginosus, *Streptococcus constellatus*, *Streptococcus difficile*,
Streptococcus dysgalactiae, *Streptococcus equi*, *Streptococcus equinus*,
Streptococcus intermedius, *Streptococcus mitis*, *Streptococcus bovis*,
10 *Streptococcus alactolyticus*, *Streptococcus galloyticus*, *Streptococcus macedonicus*, *Streptococcus infantarius*, *Streptococcus hominis*, *Granulicatella adjacens*, *Abiotrophia defectiva*, *Enterococcus avium*, *Enterococcus casselliflavus*, *Enterococcus faecalis*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus sacharolyticus*, *Gemella haemolysans* et *Gemella morbillorum*,

procédé dans lequel :

1- on met en contact un échantillon contenant ou susceptible de contenir des acides nucléiques d'au moins une telle bactérie, avec au moins une sonde d'espèce consistant dans un gène ou fragment de gène selon l'une des revendications 1 à 3, ou un oligonucléotide selon la revendication 4, et

2- on détermine la formation ou l'absence d'un complexe d'hybridation entre ladite sonde et les acides nucléiques de l'échantillon, et on détermine ainsi la présence de ladite bactérie dans l'échantillon s'il y a formation d'un complexe d'hybridation.

25 18. Procédé selon la revendication 14, caractérisé en ce qu'on cherche à détecter une espèce donnée d'une bactérie du genre *Streptococcus* et lesdits 4 genres apparentés choisis parmi les espèces :

Streptococcus mutans, *Streptococcus oralis*, *Streptococcus pneumoniae*,
Streptococcus pyogenes, *Streptococcus salivarius*, *Streptococcus sanguinis*,
30 *Streptococcus suis*, *Streptococcus acidominimus*, *Streptococcus agalactiae*,
Streptococcus anginosus, *Streptococcus constellatus*, *Streptococcus difficile*,
Streptococcus dysgalactiae, *Streptococcus equi*, *Streptococcus equinus*,
Streptococcus intermedius, *Streptococcus mitis*, *Streptococcus bovis*,

Streptococcus alactolyticus, *Streptococcus gallolyticus*, *Streptococcus macedonicus*, *Streptococcus infantarius*, *Streptococcus hominis*, *Granulicatella adjacens*, *Abiotrophia defectiva*, *Enterococcus avium*, *Enterococcus casselliflavus*, *Enterococcus faecalis*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus sacharolyticus*, *Gemella haemolysans* et *Gemella morbillorum*, procédé dans lequel, dans un échantillon contenant ou susceptible de contenir des acides nucléiques d'au moins une telle bactérie du genre *Staphylococcus*, on effectue les étapes dans lesquelles :

10 a) on réalise une réaction de séquençage d'un fragment du gène *rpoB* amplifié d'une dite bactérie donnée à l'aide des amores nucléotidiques consistant dans desdits mélanges d'oligonucléotides selon l'une des revendications 7 à 12 comprenant des séquences incluses dans la séquence SEQ.ID. n° 6 comme amorce 5' et SEQ.ID.n° 7 comme amorce 3', de préférence des séquences consistant dans lesdites séquences SEQ.ID. n° 6 et 7, et lesdites 15 séquences complémentaires, et

15 b) on détermine la présence ou l'absence de l'espèce donnée de ladite bactérie en comparant la séquence dudit fragment obtenu avec la séquence du gène complet *rpoB* de ladite bactérie ou la séquence d'un fragment du gène *rpoB* de ladite bactérie comprenant respectivement lesdites séquences SEQ ID n° 8 à 35 selon l'une des revendications 1 à 4 et séquences complémentaires, et 20 on détermine ainsi la présence de ladite bactérie dans l'échantillon si la séquence du fragment obtenue est identique à la séquence connue du gène ou du fragment de gène *rpoB* de ladite bactérie.

19 Trousse de diagnostic utile dans un procédé selon l'une des 25 revendications 14 à 18, caractérisée en ce qu'elle comprend au moins un dit oligonucléotide, mélange d'oligonucléotides, ou fragment de gène selon l'une des revendications 1 à 4 et 6 à 12.

1 2 3 4 5 6 7 8 9 10 11 12 13



FIG. 1

SEQUENCE LISTING

<110> UNIVERSITE DE LA MEDITERRANEE (Aix-Marseille II) et CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE - CNRS

UNIVERSITE DE LA MEDITERRANEE (Aix-Marseille II) et CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE - CNRS

<120> Identification moléculaire des bactéries du genre *Streptococcus*

<130> H52 437 cas 10 PCT MD

<160> 52

<170> PatentIn version 3.1

<210> 1

<211> 4523

<212> DNA

<213> *Streptococcus anginosus*

<220>

<221> misc_feature

<222> (266)..(2087)

<223> n représente a, t, c, g ou i

<220>

<221> misc_feature

<222> (266)..(4430)

<223> n représente a, t, c, g ou i

<220>

<221> misc_feature

<222> (4430)..(4503)

<223> n représente a, t, c, g ou i

<400> 1
tcatactttt agagtcaagat tttagctgctc ttttttgcc tgttttggga tttttgtcg 60
ttgtcatcaa aattaaagat tctgaaaatt actcaaaaag gataaaatgaa aattgctact 120
ctattccatt aatagagaat gtagaaaagaa gaaggagtaa aaaacttggc aggacatgaa 180
gttcaatacg ggaaacacccg tactcgctgt agttttcaa gaatcaagga agttcttgat 240
tttaccaaatt tgattgaaat ccaganggat tcgttcaaag attttcttga ccatggttt 300
aaagaagtat ttgaagatgt acttcctatc tcaaacttta cagatacaat ggagctagag 360
tttgttggtt atgaaattaa aggatctaaa tacactttag aagaagcacg tatccatgat 420
gccagctatt ctgcacctat ttttgact ttccgttga ttaataaaga aactggtaa 480
atcaaaaaccc aagaagtgtt ctttggcgat ttcccaatca tgacagaaaat gggaaactt 540

attatcaatg gtggtagcg gattatcgta tctcagctcg ttcgttctcc aggtgtttac 600
ttcaacgata aagtaracaa aaatggtaaa gttggttatg gttcaactgy cattcctaac 660
cgtggagctt ggtagagct ggaaacagac tcaaaagata ttgcttatac tcggattgac 720
cgtactcgta agattccgtt tacgacactt gttcgtgcgc ttggttttc tggcgatgat 780
gaaatctttg acatttcg cgacagcgat ctcgttcgca acacgattga aaaggatatt 840
cataaaaatc caatggattc acgtacggat gaagcgctta aagaaatcta tgaacgtctt 900
cgtccaggtg agcctaaaac agctgatagt tcacgtagtc tattggtcgc tcgtttctt 960
gatccacatc gttacgactt ggcggcagtt ggtcggtata aaatcaataa aaaattaaac 1020
attaaaaacac gtttggtaaa tcaaacgatt gcagagcctt tggtagatcc agaaacaggt 1080
gaaatcttgg ttgaagctgg aacggttatg acgcgttagt tcattgatag cattgcagaa 1140
tacttggacg gtgatttcaa taaaatcact tatattccaa atgatgcagc tgtgttaaca 1200
gagccagttg ttcttcaaaa attcaaagtg gtggcgccaa ctgatccaga tcgtgtggtg 1260
actattattg gtaatgccaa cccaggagat cgagttcata cgattacgcc agcagatatt 1320
ttggctgaga tgaattactt cttgaacctc gctgaaggac ttggcgtgt ggacgatatt 1380
gaccacttgg gaaatcgctg gattcgtgcc gttggtaat tgcttgctaa ccaagtacgt 1440
cttggcttgt ctcgtatgga gcgaaacgtt cgggagcgca tgagtgtgca agataatgaa 1500
gtgttgacac cgcaacaaat cattaacatc cgcccaagtca cagcagctat caaagaattc 1560
tttggttcat ctcaattgtc tcaattttatg gaccaacata atccactgtc tgarttgtct 1620
cacaaacgyc gtttgcgc cttggacct ggtggttga ctcgtgaycg tgctggatat 1680
gaargtgcgt gacgtgcact acacncacta tggcgtatg tgtccgattt aaacncctga 1740
vggaccaaaac atcggtttga tcaayaactt gtcttcttat ggtcatttga ataaatatgg 1800
ctttatccaa acgccgtatc gtaaagtgra tcgtgaaaca ggtctggcchaatgaaat 1860
cgtttggttg acagcggang aagaagatga atttattgtt ggcgtttttt ggcgtttttt 1920
aacagaagat ggtcggtttt cagaagcgat tgtcatggga cgtcaccaag ggaacaacca 1980
agaatttcct tcagatcarg trgatttcat ggatgtgtcg cctaaaggagg tagttgccgt 2040
tgcgacagca tgtantccnk ttccytgaaa aygnacgact caarccntgn tstcatgggt 2100
gccaaacatgc aacgtcaagc sgtaccgttg attgatccgc atgcaccata ygywggtana 2160
tggatggaa taccaagcag antsaygamt ctggcggc tgattantgc mcaacacgac 2220
ggtaaagttg tmtattygtga tgcagccaaa gttgaagttc gtcgtgaaga tggctcactt 2280

gtatgtntat catagntgac gaaattccgc cgtnaaact gstdggtaacgt tgmttacaac	2340
acaacgttagc ggstdggtaaa agattggcga tacagntgta aaaaggtgta stttatcgca	2400
gacggacctt ctatggaaaa aggtgaaatg gcrcttggac aaaayccaat cgttgcttat	2460
atgacatggg aaggttacaa ctttgaagat gccgttatca tgagtgagcg httagtgaaa	2520
gacgatgttt acacatctgt tcacttggag gaattcgaat cagaaacacg tgatacwaag	2580
cttaggmccct gaagaaatca ckcgcgaaat tccaaacgty ggtgaagatg ccnttygasa	2640
gaccttggac gaaaygggra ttataccgya ttggtgcyga rgttaaagag ggcgacattc	2700
ttgttgtaa agtcacacca aaaggtaaaa aagatcttc tgctgaagag cgtctttgc	2760
acgcaatctt cggtgacaag tcacgtgaag tacgtgatac ytcycttcgt gtaccwcayg	2820
gtgsygcatt gkgyygtycg tgatgtgaaa atcttwactc gtgcsaacgg tgatgaattg	2880
caatcwggtg tcaacatgtt ggtacgtgtt wcacntcgct caaaaacgka araycamgtg	2940
tyggrgataa gatggcyggw cgtcacggaa acaaagggtt tgttccgc attgttccag	3000
ttgaggatat gccgtatctt ccagatggaa caccagttga tattatgttg aaccacttg	3060
gggtgccatc tcgtatgaat attggtaag ttatggagct tcacctcggt atggctgctc	3120
gcaaccttgg cattcacatt gcaacaccag tatttgacgg ggctagctca gatgatctt	3180
gggaaaccgt tcgtgaagct ggcattggata gcgtatgctaa gacaatcatt tatgatggcc	3240
gtactggta gccatttgat aatcggttat ccgttgggt catgtacatg atcaaactcc	3300
accatatggt tgatgataag ctccatgccc gttccgttgg tccttattca accgttacgc	3360
aacaacctct tggtgtaaa gcgcagtttg gtggacaacg ttttggagaa atggaagttt	3420
gggctttga agcctacggc gcttctaact tccttcaaga aatcttact tacaagtcag	3480
atgacatcaa tggtcgtttg agagcttatg aagccattac caaaggtaag ccaattccaa	3540
aaccaggtgt tccagaatcc ttccgtgtcc ttgtaaaaga attgcaatca cttggctttg	3600
acatgcgtgt ctttgcgtt gacgacaatg aagtcgaact tcgtgacttg gacgaaggca	3660
tggatgatga tgtgattcat gtagacgatc ttgaaaaagc acgtaaaaaa gcagcacaag	3720
aagcaaaagc cgcttttgat gctgaaggaa aagaataaga actgattcaa tagataataa	3780
agaaaggtaa gaaatagtgg ttgatgtaaa tcgtttcaa agtatgcaaa tcaccctagc	3840
ttctcctagt aaagtccgct cttggcttta tggagaagtg aagaaacctg aaacaattaa	3900
ctaccgcaca ctaaaaaccag aacgcgaagg gcttttgat gaagtcatct ttggcctac	3960
gaaagactgg gaatgtgcgt gtggaaaata taaacggatt cgttataaaag gaatcatttg	4020

tgaccgttgt	ggtgttgaag	taactcgta	taaagtttgt	cgtgaacgt	tggacat	4080
tgagttgaaa	gccccagtct	cctcatattt	ggtattttaa	aggaattcca	antcgcatgg	4140
gcntgacctt	ggacatgagc	cctcggtctc	ttgaagaagt	catntanttt	gcagcttatg	4200
tggtgantga	ccctaaagat	acnccacttg	agcacaaatc	cattatgaca	gagcggatg	4260
gttngtgaac	gctgacntga	atatggccaa	ggctctttt	ttgcaaaaat	gggtgytgaa	4320
gcaatccaag	atctnnntgaa	acangtagac	ntggaaaaag	aaattgcaga	gctcaaagat	4380
gaattaaaaaa	cggcaagtgg	gcaaaagcgc	gtaaamgcta	anttcgtcgn	tnngactctt	4440
ttcgatnctt	tccaaaaatc	atggtacaca	aaaccagaac	tggatggtct	taaaccatcn	4500
ntntcaccgc	tcattccaga	cac				4523

<210> 2

<211> 4118

<212> DNA

<213> Streptococcus equinus

<400> 2

cacgcgttgt	cgacggcccg	ggctgggtgaa	ttgtcataag	ttgtgttagta	gtaaattccc	60
ttatcagtgt	tgtatgcata	gctataaata	gtgtactcat	atttgccact	ttcatcgaca	120
tagcaaagtc	ctttttgttg	ttcaacggat	tttaaaatgt	ggaagaattg	attaacactg	180
ctttcttctg	tttcttcagc	cacagaattt	aatttgtaa	aagtaacttt	tacataacgt	240
gacattgatg	ataaaatcacc	aggcaagcca	agtccaccca	tgccacggct	ataagttca	300
agttctaact	cttttagcaaa	acgattttct	gaaacctttg	gagatagatg	acgatagttt	360
ttcaaaattga	ataattgttt	atcaaaaatgtt	ggattattag	tcaaaacacc	tgttgagttt	420
ttcgtaaact	tatagggcac	gcgtggtcga	cggcccgccc	tggtaaagac	ttcttgata	480
acggattaaam	agaagttttt	gaagatgtac	ttccgattac	aaactttacg	gatactatgg	540
agcttgaatt	tgttggttac	gaattgaaag	agcctaagta	tacgcttgaa	gaagctcgta	600
tccacgatgc	atcttattca	gcacctattt	ttgttaacctt	ccgtttgatt	aataaaagaaa	660
caggagaaat	caaaactcaa	gaagttttct	tcggtgattt	cccaattatg	actgaaatgg	720
gtacattcat	catcaacggt	ggtgaacgt	ttatcgtttc	tcagttggtt	cgttctcctg	780
gtgtttattt	caacgataaa	gttgataaaa	acggtaaagt	tggttacggt	tcaactgtaa	840
tccctaaccg	tggagcatgg	cttgaattag	aaacagattc	aaaagatatt	gcttacacac	900
gtatcgaccg	tacacgtaaa	attccattta	caactttgt	acgtgcgtt	ggtttctcag	960

atacatctgt tcacttggaa gaatacgaat cagaaacacg tgatactaag ttaggcctg 2760
 aagaaatcac tcgcgaaatt ccaaacgttgcgtgaagatgc cttcgcaac ttggacgaaa 2820
 tgggattat ccgtatttgtt gccaaggta aagagggcga cattttgtt gtaaagtca 2880
 caccaaaagg tgaaaaagat cttctgttg aagagcgtct cttgcacgca atcttcggtg 2940
 acaagtcacg tgaagtacgt gatacctctc ttctgttacc tcacggtgcc gatgggtcg 3000
 ttctgtatgt gaaaatcttt actcgtgcca acggtgatga attgcaatca ggtgttaaca 3060
 tgttggttcg tgttcacat cgctcaaaaaa cgtaagatca aggtcggaga taagatggcc 3120
 ggtcgccac ggttaacaagg gtgtcggttc acgtaywgta cctgttgagg atatgccata 3180
 tcttccagat ggaacaccagg ytgacawcat gttgaaccca ctsggggtgc catwcgtat 3240
 gaacatcgga caagttatgg agcttcaccc ttgtatggct gctcgtaacc ttgttattca 3300
 cattgcaaca ccagtcttg atggggcaac ttctgaagac ctttggata cagttaacga 3360
 agctggtatg gctagcgacg ctaagacagt tcttacgat ggacgtactg gtgaaccatt 3420
 tgataaccgt gtgtcagttg gtgtcatgta catgattaaa cttcaccaca tgggtatga 3480
 taaacttcac gcacgttcag ttggcctta ctcacttgtt acgcaacaac ctcttggtg 3540
 taaagcacaa ttgggtggac aacgttcgg tgaaatggaa gtttggctt tggaagctta 3600
 cggtgcata aatgttcttc aagaaatctt gacttacaaa tcagatgttgc tcaacggcgt 3660
 tcttaaagct tatgaagcca tcactaaagg taaaccaatt ccaaaaccagg gtgtccaga 3720
 atcattccga gttcttgtaa aagaattgca atcacttgtt cttgacatgc gcgtgcttga 3780
 tgaagatgac aatgaagtag aacttcgtga tcttgcgtt ggtgaagatg acgtgttat 3840
 gcacggttgcgtt gatcttgcgtt aagctcgta aaaacaagaa gcagaagaag cggaaaaagg 3900
 agaagtttct gcagaagaaa acaaataata gaaagaaca ttccagacatg agagaggcaa 3960
 gacctgcttc tcttggtcag attgtttgtt tgagtccat aacgataat gatgtcttac 4020
 gaatcatgaa ttgttaagtc atgacagtta gaaagtagcg cagctatttc aaagtcttac 4080
 gaaggtatca tggtgacgta atcggttacag ccggcgtc 4118

<210> 3

<211> 3425

<212> DNA

<213> *Abiotrophia defectiva*

<400> 3

atataggca cgcgtggcgtc acggcccggtt ctggcctaa acaacatgta acgtcactcc 60
 gatgagttgg ttctgttgc tttttttgc gttcaaaga ccgaaaaatg tcattgtca 120

acaattatta ataattgtaa ccttaatgt aagtgggtt ctttagattat attatagggg 180
tgaatcgctt gagtcatatac gtgaaatacg gtaaaaaaagc tgagcgtcga agctatgcgc 240 ..
gtatcgacga agtcttagag ttgccgaact tgattgaaat ccaaacggat tcctacaaat 300
ggttcttgg a tgaaggccta aaagtgtat tcgaggacat ttgcggatt gtcgaccatt 360
cggagaactt ggaacttcat ttttagact atgagttcaa ggaagctaag tatacgcttag 420
aagaagctcg tagccatgac gctaactact caaaaccaat ctatgttaacc ttgcgcctgt 480
tcaacaaaga gacaggtgaa gtcaaagaac aagaagtctt cttcgggac ttccaaatca 540
tgaccgaaat ggggaccttc attatcaacg gggcggaacg gttatcggtt tccagttgg 600
tacgttctcc aggtgtctac ttccacgacc gtatggacaa gaaaggccgc cacagctata 660
cttctacggt tattcctaac cgtggggctt gtttggatt tgaatcagat gctaagggga 720
ttgcctacgt ccgcattgac cggacccgga agattccatt gactgtctt atgcgtgcct 780
taggtttgg ttcatgtac gagatttatg atatcttcgg ccaatctgag ctcttagact 840
taactatcga gaaggatgtt cacaaaaaca ttcaagactc tcgtacggaa gaagccttga 900
aggacattta cgagcgtctc cgtccaggtg aacctaagac cgcagaaagc tcacgttaacc 960
tcttgggtgc gcgcttcttc gacccacgtc gctatgactt agcacctgta ggtcgttata 1020
agatcaataa aaagctccac ctcaagaacc gtttgggtgg cttgactttg gctgaaacct 1080
tggtaaccc agaaacaggc gaagtgcctt ttgaagaagg aacggtctt gatcaagaac 1140
gtgttcaagc cctgattcca tacttagagg ctggcttgaa taaggtaacc ctctatcctt 1200
ctgaagatag tgtggtagct caaccaattt atttacaaat catcaaagtt tattcaccta 1260
agaacgcccga gcaagtgatt aacatcatcg gtaacggaa cattgagaag attaagtgt 1320
tgacgccagc tgacattatt gcgtcaatga actactatct ctatggac caaggaattt 1380
gtgtgacaga tgatatcgac cacttggcta accgtcgat tcgtttagtc ggtgaattat 1440
tgcaaaacca attccgtatc gggctatccc ggatggaaacg ggttagtgcgt gaacgtatgt 1500
cgctccaaga ttttgcgacc atcacaccgc aacaatttgc taacattcgat ccagtagtgg 1560
cggttattaa ggaattcttc gttcatccc agttgtcaca attcatggac caagtttacc 1620
cactcgggaa attgacccac aaacgtcgatc tgcgttcattt agggcctggat ggtttgacgc 1680
gggaccgtgc cggctatgaa gtgcggacg ttcaactactc tcactacggc cgtatgtgtc 1740
caatcgagac gccagaaggt cctaacatcg gtttggattaa cagcttgcgt tcttacgtt 1800
agattaacaa gtatggttt attgagacgc cttaccgtaa agtggacaaa tcgggttacgc 1860

cacaccgtgt cacgaccgaa attgactacc tagcagcgga cgaggaagac ttgtacgtag	1920
tagcccaagc caactctaaa ctcaacgaag acgggacctt cgccaatgac ctagttatgg	1980
cgcgtttccg ttcacaaaaac attgaggtta acgttgacca agtagactac atggacgtat	2040
cgcacaaaaca ggttgcgct gtcgcgactg ctagcattcc gttcttgaa aacgacgact	2100
ccaaccgggg cttgatgggt gccaacatgc aacgtcaagc tgtgccactt attaatccac	2160
aatccccact gattgggact gggatggaat ataaggcagc acacgactct gggctgcgc	2220
tcttatgtaa gcgcgcgggt gaagtggttt atgtcgatgc taacaaggtg cgctgcgca	2280
ctccagaagg tgaagttgac gaataccgtt taaccaagtt tgcacgttct aacgctggga	2340
cctgttacaa ccaacgtcca atcgtagaat taggcgacca agttgatgcc ttggaaatct	2400
tagcagatgg tccatctatg caaaatgggg agatggccct cggtaaaaac ccactggtag	2460
ccttcatgac ttgggaaggg tataactatg aggacgcgggt tatcatgtct gaacgtctgg	2520
tcaaagacga ttttatacc tctatccaca ttgaagaata tgaatcagag tccctgtaya	2580
cyaagttagg ccctgaagaa attacacgcg aaattccaaa cgtgtccgaa gatgccctca	2640
agtacttaga caaagacggg attatctgta tcggggcgga agtaaaagac ggcgatatct	2700
tagttggtaa ggtaaacacca aaagggtgtga ccgagttgtc tgcggaaagaa cgcttgctcc	2760
atgctatctt cggtgagaag gcgcgtgaag tacgtgatac ttccctgcgt gtgccacacg	2820
gcggggggcggtt gattgtccac gacgttaaaa tcttacccg cgaagctggc gacgaattgg	2880
caccaggtgt caacaagcta gtccgcgtct acatcgtaaaa acgttaaaa atcaatgaag	2940
gggataagat ggccggcgtt cacggtaaca aagggttgtt ctcccttatac atgcccggaa	3000
aagatatgcc attcttacca gatggtaccc cagttgatcatgttgcac ccattagggg	3060
ttccatcccg tatgaacatc gggcaagtcc tagagttaca cttggggatg gctgctcg	3120
aaatgggcattt caagattgca acacctgtct ttgacgggtgc tagtgaagaa gatgtctgg	3180
aaacagttaa ggaagccggc tttagaagctg acgctaagac tatcttatcatgttgcggaa	3240
ccggtaacc atttgaccgt aaagtctctg ttggggttat gtacatgatt aagttggccc	3300
acatggtcga tgacaagttg cacgcccgtt caacaggtcc atactctctg gttacccaac	3360
aaccattggg tggtaaagct caatttggtg ggcaacgttt cggggagatg gaggtttggg	3420
cccta	3425

<212> DNA

<213> Streptococcus mutans

<220>

<221> misc_feature

<222> (619)..(3193)

<223> n représente a, t, c, g ou i

<400> 4

ggaccctttt atgacttctt ggatacaggt ctgaaggaag ttttgaaga tgtgcttcca 60
atttccaatt tcacagacac tatgaaatta gagtttgtgg gttatgagtt gaaagagcct 120
aagtatacat tggaagaagc acgtgctcat gatgcacatt attctgcccc catctttgtt 180
actttccgtc tcatcaataa agaaaactggt gaaattaaga cacaagaagt atttttggt 240
gattttccct tcatgactga aatgggtact tttattatta atggctgctga acgtattatc 300
gtttctcagt tggtacgttc accaggtgtt tatttaatg ataaagtggta taaaaatggg 360
aaaattggct atggttcaac ttttatccct aaccgcggtg cttggcttga gcttgaacg 420
gactctaagg atattgctta tactcgtatt gatcgtaatc gtaaaattcc ttttacgacg 480
ctggttcgtg cactcggtt ttccggggat gatgagatta ttgatatttt tggtgatagc 540
gaattggttc gtaataccat tgaaaaagat atccataaaa atcctaataatgc ctctcgtaca 600
gatgaagctc tcaaggaant tatgaacgtc ttcgtccggg tgaacctaaa acggcagatt 660
cntcacgcag ttttctgatt gcacgtttct ttgatgcgcg ccgttatgat tagcagctgt 720
tggccgctat agataataag aagttaaacg tcaaaaacggg tctttgaatc aagtcattgg 780
ctgaaaanna gtagatctga aacaggcgaa attcttggta aaagctggta ctgaaatgac 840
acgcagtgtt attgattcga ttgcagatta tcttgatgga gatctcaata aaatgttt 900
tacgccaaat gaatacgctg ttttgacaga acctgttgc tttcaaaaat tcaaaatgtt 960
ggctccaaat gatccagacc gcacggttac ttttattttt aatgccagtc caagatgaca 1020
aagtacgtca cttgacacca gccgatacgt attagctgaa atgtcttatt tccttaactt 1080
ggctgagggt ntaggtaaag ttgatgatat tgaccattta ggcaaccgac gtattcgtgc 1140
ttttgggttta ttgcttgcta atcaatttcg tattggtttgc acgttatgg aacgcaatgt 1200
tcgtgaacgc atgtccgttc aagataatga agtcttaacg ccacaacaga ttatataat 1260
tcgcccgtta acagcggcaa ttaaagagtt ttttggttct tctcaattgt cacagttcat 1320
ggaccaacac aatccactgt ctgaattgtc tcataaacgc cgtttgcag ctttaggtcc 1380
ttgtggttta acacgcgacc gtgctggta tgaagtccgt gatgtgcact atacgcatta 1440

tggtcgatg tgtccaaattg aaacgcctga aggacccaaat attggattga ttaataactt 1500
gtcttcctat ggtcatctta ataaatatgg atttatccaa acaccatacc gtaaagttga 1560
ccgtgagaca ggtaaagtaa ccaatgaaat cgaatggctt actgctgatg aagaagatga 1620
attcactgta gctcaggcta actcaaaaact caatgaagat ggaagctttg ctgaagaaat 1680
cgtcatggga cgtcatcaag ggaataacca agagttcca gcaagttctg ttgaatataat 1740
ggatgtttct cctaagcagg tagttgcggt agcgacagca tgtattcctt tccttgaaaa 1800
tgatgactcc aaccgtgccc ttatggagc taacatgcag cgccaaagctg tgccattgat 1860
tgatcctaaa gcaccttttgg tatggaaacttgg tatggaaatat caagcagcccc atgattctgg 1920
agccgctatt atcgctcaac ataatggaa agtggtttat tccgatgcag ataagattga 1980
agttcgccgt gaagatggct cactagatgt ttatcatgtt accaaattcc gtcgttctaa 2040
ctctggaaact gcctacaatc aacgtactct tgtagggta ggcgatagtg ttgagaaggg 2100
ggactttatt gcagatggtc cttctatgga aaagggtgag atggctcttgc gacaaaatcc 2160
agtggttgct tacatgactt gggagggta caactttgaa gatgctgtt tcattgagcga 2220
gcgtcttgc aaggatgatg tttatacttc tgtccatttta gaagaatttg aatctgaaac 2280
tcgtgataca aagcttggac ctgaagaaat tacgcgtgaa atcccaaatttgc ttggtaaga 2340
tgccctgaaa gaccttgcatttgc aaatggaaat tattcgatttgc ggtgctgagg ttaaaagaagg 2400
tgatattcta gttggtaaaat tgactcctaa aggagaaaaaa gatctttctg cagaagaacgc 2460
cctcttgcatttgc gtcatttttgc gtgacaaatc acgtgaagtt cgtgatactt ctcttcgtgt 2520
acctcatggc ggcgacggcgtg ttgtttgtga tgtgaaaatc ttacacgtg ctaatggaga 2580
tgaacttcaa tcaggtgttta acatgctggc tcgtgtttat atcgctcaaa aacgtaaaaat 2640
caaggtcgga gataagatgg ccggacgtca tggtaacaag ggtgtcgaaaatccgttattgt 2700
accagtggaa gatatgccat atcttccaga tggaaacacact gttgatatca tgcttaatcc 2760
acttgggtgc ccatcacggc tgaacattgg gcaagttatg gaactccatc ttggatggc 2820
tgctcgtaat ttggcatttc atattgcaac gcctgtctt gacggagcaa cttctgatga 2880
tctttggaa acagtaaaatg aagccggat ggtttctgat gctaaaactg ttctttatgt 2940
tggtcgcaca ggggagccgt ttgataatcg tgtatcagtt ggtgttatgt atatgattaa 3000
acttcaccac atggttgatg ayaaccattt tgtctatgca magwtcagtt ggcccttakt 3060
caaygawtam tcagasgart tcctgctwgg tgtaaaggct ncaattgtct ttagaggtta 3120
aggctggta aataacggta tgctggattt gatggcaatg ggcaagtgaa tantcaacac 3180

cggccgtcta cancgtgc

3198

<210> 5
<211> 3096
<212> DNA
<213> Enterococcus faecalis

<400> 5
gacccttatac aattggtttt tagatgaggg acttcgtgaa atgtttgaag acattttacc 60
aattgatgat ttccaaggaa acttatacctt agaatttgtt gactatgaat taaaagaacc 120
aaagtacaca gtagaagaag cccgcgcaca ttagtccaaac tattctgcgc cattacatgt 180
aacattacgt ttaaccaacc gtgaaacagg tgaaattaaa tcccaagaag tcttcttcgg 240
cgatttccca ttaatgacag aaatgggtac cttcatcatc aacggggcag aacgtgttat 300
cgtttcccaa ttagttcggtt ctccaggtgt ttacttccat ggaaaagtgg acaaaaacgg 360
caaagaaggt tttggctcaa cagtcattcc taaccgtggt gcatggtag aaatggaaac 420
agatgcgaaa gacatttctt atgttcggat tgaccgcaca cgtaaaattc cttaactgt 480
gttagttcgt gcttttaggtt tcgggtcaga tgataccatc ttgaaaattt tcggcgacag 540
cgaaagctta cgcaacaccaa ttgaaaaaga tttacacaaa aacgcaagtg attctcgtac 600
agaagaaggc ttgaaagaca tttatgaacg tcttcgccc ggcgaaccaa aaacagcaga 660
tagtcacgt agcttgttaa cttgcacgtt tctttgatcc aaaacgttat gattggcaa 720
acgttggtcg ctacaaagtt aacaaaaat tagacttaa aacacgtcta tttaacttaa 780
ccttagctga aacgctagtt gatccagaaa ctgggtgtaaa tcattgtcga aaaaggcaca 840
gttttaacac actacatcat ggaaacatta aggcrataca ttgacaaacg gcttaaacag 900
cgtaacttac tatccaagtg aagatgcggt agtaactgaa ccaatgacga tccaaatgtat 960
tcaagttctt tcacaaaaag atcctgaacg tatcgtaaat gtgattggta acggctatcc 1020
agacgacagc gtaaaaacag ttctgtccagc agatatcggtt gcttcaatga gctacttctt 1080
caacttaatg gaagatatcg gtaatgtcga tgacatcgac cacttaggtt atcgctgtat 1140
ccgttcagta ggcgaattat tacaaaacca attccgtatt ggtttagccc gtatggacg 1200
tgtggttcgt gaaagaatgt ctattcaaga cacagaaaca ttgacaccac aacaattaat 1260
taacatccgt ccagtggttag caagtatcaa agaattctt ggttcttcac agttatcaca 1320
gttcatggac caaacaacc cattaggtga gttaaccat aaacgtcgac tatcagcctt 1380
agggcctgggt ggtttgactc gtgatcgac cggttatgaa gttcgtgacg ttcactactc 1440
tcactatggt cgtatgtgtc caattgaaac gcctgaggga ccaaataatcg ggttcatcaa 1500

tagcttatct	agttatgcga	aagtgaataa	atttggtttc	atcgaaacgc	cttatacgccg	1560
tgttgatcgt	gcgacaggcc	gtgttactga	tcaagtagat	tacttaacag	cagacatcga	1620
agaccattat	atcgtagcgc	aagcgaactc	acttttaaat	gaagatggca	catttgccaa	1680
tgatgttgtt	atggcgcgtc	tacaaagtga	aaacttagaa	gttgcgttag	acaaaagttga	1740
ctacatggac	gtttcaccaa	aacaagtgt	cgcagtcgca	acagcatgta	ttcccttctt	1800
agaaaaacgat	gactccaacc	gtgcctttagt	gggtgcacac	atgcagcgtc	aagcggtgcc	1860
gttaattcaa	ccacgctctc	cgtgggtagg	tacaggtatg	gaatataaat	cagcccatga	1920
ctcaggtgct	gttttactat	gtaaacatga	cggtgtcgta	gaattcgtcg	atgcaaaaga	1980
aattcgcgtt	cgtcgcgaca	atggcgcatt	agacaaatat	atggttacaa	aattccgtcg	2040
ttcttaactca	ggaacaagct	acaaccaacg	cccaattgtt	cacttaggtg	aaaagttgaa	2100
aaggcgatac	tttaccggat	ggaccttcta	tggaagaagc	gaaatggctt	tatggcaaaa	2160
cgtcttagtt	gccttcatga	catgggaagg	ttacaactac	gaggatgcca	ttatcatgag	2220
ccgtcgttta	gttaaagacg	atgtotacac	ttctgtgcat	attgaagaat	atgaatcaga	2280
agcacgtgat	acaaaattag	gacctgaaga	aattaccgt	gaaattccaa	acgttgggga	2340
agacgcgttg	aaagacttag	acgaaatggg	gattatccgc	attggtgctg	aagttcaaga	2400
tggcgactta	ctagttggga	aagtcacacc	taaaggggtc	acagaattat	ctgcagaaga	2460
acgtttatta	cacgcaatct	tcggggaaaa	agcccgcgaa	gttcgtgata	cgtctctccg	2520
tgtacctcac	ggtggcggcg	gtatcgttca	tgatgtgaaa	atcttactc	gtgaagctgg	2580
cgtatgaattta	tcaccaggtg	tcaacatgtt	agttcgtgtc	tatatcggtc	aaaaacgtaa	2640
aattcacgaa	ggagataaaaa	tggcgggacg	tcacggaaat	aaaggggttg	tttcccgat	2700
tatgccggaa	gaagatatgc	cattcttacc	tgacggaaaca	cctgttgata	tcatgttgaa	2760
cccattaggg	gtaccttctc	gtatgaatat	cggacaagta	cttgaattac	acttaggtat	2820
ggctgctcgc	caattaggtt	ttcacgtcgc	aacacctgtt	ttcgatgggg	caaccgatga	2880
agacgtttgg	gaaactgttc	gtgaagctgg	tatggcttagc	gatgctaaaa	cagttcttta	2940
cgtatggacgt	acaggtgaac	catttgataa	ccgtatttcc	gttgggtgtca	tgtatatgat	3000
taaattagcc	cacatgggtt	atgacaaatt	gcatgctcgt	tcaatcgac	cttactctct	3060
tgttaacgcaa	caaccgttgg	gtgtaaagct	caattc			3096

<212> DNA
<213> amorce

<220>
<221> misc_feature
<222> (6)..(6)
<223> n représente a, t, c ou g ou i

<400> 6
aarytnngmc ctgaagaaat

20

<210> 7
<211> 23
<212> DNA
<213> amorce

<220>
<221> misc_feature
<222> (3)..(3)
<223> n représente i

<220>
<221> misc_feature
<222> (3)..(3)
<223> n représente a, t, c ou g ou i

<400> 7
tgnartttrt catcaaccat gtg

23

<210> 8
<211> 709
<212> DNA
<213> Streptococcus suis

<400> 8
cgcgaaattc caaacgttgg tgaagatgcc cttcgcaact tggacgaaat ggggattatc 60
cgtattggtg ccgaagttaa agagggcgac attcttggtg gttaagtcac accaaaaggt 120
aaaaaaagatc tttctgttga agagcgtctc ttgcacgcaa tcttcgggtga caagtcacgt 180
gaagtacgtg atacctctct tcgtgtacct cacggtgccg atgggtgtcgt tcgtgtatgt 240
aaaatcttta ctctgtccaa cgggtatgaa ttgcaatcag gtgttaacat gttggttcgt 300
gtttacatcg ctcaaaaacg taagatcaag gtcggagata agatggccgg tcgtcacgg 360
aacaagggtg tcgtttcacg tattgtacct gttgaggata tgccatatct tccagatgga 420
acaccagttg acatcatgtt gaaccactc ggggtgccat cacgtatgaa catcggtcag 480
gttatggaac ttcacttggg tatggcggct cgcaacttgg gcatccatat cgcaacacca 540

14/32

gttttcgatg gtgcaagttc agaagacctc tggtaactg ttaaagaagc aggtatggac 600
 tcagatgcca agaccattct ttacgatgga cgtacaggtg aaccatttga caaccgtgta 660
 tctgttggtg tcatgtacat gatcaagctt caccacatgg ttgatgaca 709

<210> 9
 <211> 725
 <212> DNA
 <213> *Streptococcus sanguinis*

<400> 9
 tgcgtatcaac catgtggta gcttaatcat gtacatgaca ccgacagata cacggttgtc 60
 aaacggctca ccggtaacgtc catcgtaaag aatagtcttg gcatcgctat ccataccagg 120
 ttcacggaca gtatcccaga ggtcttctga gcttgctcca tcaaagaccg gtgtcgcaat 180
 atggatgccc aagttacgtg ctgccataacc aaggtgaagc tccataacct gaccaatgtt 240
 catacgtat ggtaccccgaa gtgggttcag catgatataca actgggtttc cgtctggcaa 300
 ataaggcatg tcttccacag gaacgatacg ggatacaacc cccttggttc cgtgacgacc 360
 agccatctta tctccgaccc tgcgttacg ttttgagcg atgttagacac gaaccaacat 420
 attaacgcca gattgcaact catcaccatt agcacggta aagatcttca cgtcacgaac 480
 cactccatca gcaccgtgcg gcacacgcag agaggtatca cggacttcac gagacttgc 540
 tccgaagata gcgtgcaaga ggcgcttttc agcagaaaga tcttttccac ccttaggggt 600
 aactttaccc acaaggatata cgccttcctt gacttccgcc ccgatgcggta taatacccat 660
 ttgcgtccaaa ttgcgttaggg catcttcccc tacgtttgga atttcgcggg taattcttca 720
 ggtca : 725

<210> 10
 <211> 728
 <212> DNA
 <213> *Streptococcus salivarius*

<400> 10
 ttgcgtatcaa ccatgtgtga agtttgcgtca tgtacatgac accaactgat acacggttat 60
 caaatggttc acctgtacgt ccatcgtaaa ggattgtctt agcatcacta tccataccctg 120
 cttaacgaac agtatcccag aggtcttctg agcttgcacc gtcaaagact ggtgttgcga 180
 tgtggatacc caagttacga gcagccatac caaggtgaag ttccataacc tgaccgtgt 240
 tcatacgtga tggcacccca agagggttca acatgatatac aactggtgta ccgtctggaa 300
 ggtaaggcat gtcttcaaca ggaacaatac gagaacaacaac cccttgcgtta ccgtgacgac 360

15/32

cgcccatctt atctccgacc ttaatcttac gttttgagc gatgtaaaca cgaacaagca	420
tgttaacacc tgattgcaat tcatcaccgt ttgcacgtgt gaagattttt acatcacgaa	480
cgacaccatc accaccgtga ggtacacgga gtgaggtatc acgtacttca cgagatttat	540
caccaaagat agcatggaga agacgttctt cagcagaaag gtcttttca cccttaggtg	600
ttaccttacc aacaagaatg tcacccttctt taacctcagc accgatacgg ataataccca	660
tttcgtcaag gtctttgaga gcttcttcac caacgtttgg caattcacgt gtaatttctt	720
caggtcca	728

<210> 11
 <211> 725
 <212> DNA
 <213> *Streptococcus pyogenes*

<400> 11	
tgtcatcaac catgtggta agtttgcata tatacatgac accaacggat acacgggtgt	60
caaatggttc accgggtgcga ccatcataaa ggaccgtctt agcatcgcta tccataccag	120
cttcacgaac agtgccttccaa aggtcttctg atgaagcccc gtcaaagaca ggtgttgcaa	180
tgtgaataacc aagattacga gcagccatac caaggtgaag ttccataacc tgaccaatat	240
tcatccgtga tggcacccca agagggttca acatgatgtc aactgggttt ccgtctggaa	300
ggtatggcat gtcttcaact ggtacaatac gtgaaacgac acccttgttt ccgtgacgac	360
cggccatttt atctccgacc ttgattttac gttttgagc gatgtaaaca cgcacaagca	420
tattaacacc tgattgcaat tcatcgccgt tagcggtgt aaagattttc acatcacgaa	480
cgataccatc accaccgtga gggacacgaa gtgaggtatc acgcacttca cgcgatttat	540
ccccaaagat ggcgtgaagt aaacgttctt cagcagaaag gtcttttca ccttaggtg	600
tgactttacc tactaagatg tcgccttctt taacctcagc accgatacgg ataatgcccc	660
tttcgtcaag gtctttgagg gcttcttcac caacatttgg gatttccgag tgattcttca	720
gggca	725

<210> 12
 <211> 724
 <212> DNA
 <213> *Streptococcus pneumoniae*

<400> 12	
caaccatgtg gtggagtttgc atcatgtaca tgactccgac agaaaacacg gttatcaaac	60
ggttcaccag tacgtccatc gtaaaggatc gttttggcat cgctatccat acctgcttct	120

16/32

ttaacagttg accaaagatc ttcagaactt gctccatcaa agactggtgt cgcgatgtga	180
ataccaagag tacgagctgc cataccaagg tgaagctcca taacctgacc gatattcata	240
cgtgatggta ccccaagtgg gttcaacatg atgtcgactg gagttccgtc tggaaggtaa	300
ggcatgtctt ctacaggaac gatacgagag acaacccctt tgttccgtg acgtccggcc	360
attttatctc cgacccttaat cttacgttt tgagcgatgt aaacacgaac caacatgtta	420
acacctgatt gcaactcatc tccatttaca cgtgtaaaga tcttaacatc acgaacgaca	480
ccatcggcac cgtgtggtagt acgaagagaa gtatcacgca cttcacgaga cttgtctcca	540
aagatagcgt gcaagagacg ttcttcagct gaaagatctt tctcacccctt aggtgttact	600
ttacctacaa gaatatcacc ttcttaacc tcagcaccaa tacggataat cccatttcgt	660
caaggtcttt gagggcatct tcaccaacgt tttgaaattt cgcgagtgtat ttcttcaggt	720
ccaa	724

<210> 13

<211> 694

<212> DNA

<213> *Streptococcus oralis*

<400> 13

actcgtaaaa ttccaaacgt tggtaagat gcccctaaag accttgacga aatgggtatt	60
atccgtattt gtgctgaggt taaagaagga gatatccttgc taggtaaagt cacacctaag	120
ggtaaaaaag acctttctgc tgaagaacgt ctcttgcacg ctatcttcgg agacaagtct	180
cgtgaagtgc gtgataacttc tcttcgagta cctcacgggtc ccgatgggtgt cgttcgtgtat	240
gttaagatct ttacacgtgc aaatgggtat gagttgcaat ctgggtgtaa tatgctggtt	300
cgtgtctaca tcgctcaaaa acgtaagatc aagtccggaga taagatggcc ggacgtcacg	360
gaaacaaaagg ggttgcgtct cgtatcggtt ctgttagaaga catgccttac cttccagatg	420
gaactccagt cgatatcatg ttgaacccac ttgggggtgcc atcacgtatg aatatcggtc	480
aggttatgga actccacccctt ggtatggcag cccgtactct tggatccac atcgcaacac	540
cagtctttga cggagcaagt tcggaagacc tttggacac tggtaaagaa gcaggtatgg	600
atagcgatgc caaaaacaatc cttaacgtatg gacgtacagg tgagccgttt gacaaccgtg	660
tatcagttgg tgtcatgtac atgatcaaac tcca	694

<210> 14

<211> 728

<212> DNA

<213> *Streptococcus mutans*

<400> 14

tgtcatcaac catgtggta agttaatca tatacataac accaactgat acacgattat 60
 caaacggctc ccctgtgcga ccatcataaa gaacagttt agcatcagaa tccataccgg 120
 cttctttac tgttcccaa agatcatcag aagttgctcc gtcaaagaca ggcgttgcaa 180
 tatgaatgcc caaattacga gcagccatac caagatggag ttccataact tgcccaatgt 240
 tcatccgtga tggcacccca agtggattaa gcatgatatc aacaggtgtt ccatctggaa 300
 gatatggcat atcttccact ggtacaatac gggaaacgac acccttgtta ccatgacgtc 360
 cggccatctt atctccgacc ttgattttac gttttgagc gatataaaca cgaaccagca 420
 tgttaacacc tgattgaagt tcatctccat tagcacgtgt aaagatttc acatcacaaa 480
 caacaccgtc gccaccatga ggtacacgaa gagaagtatc acgaacttca cgtgatttgt 540
 caccaaaaat ggcatgcaag aggcggttctt ctgcagaaag atcttttctt cctttaggag 600
 tcactttacc aactagaata tcaccttctt taacctcagc accaatgcga ataattccca 660
 tttcatcaag gtcttcagg gcatcttcac caacatttg gatttcacgc gtaatttctt 720
 caggtcca 728

<210> 15

<211> 730

<212> DNA

<213> *Streptococcus mitis*

<400> 15

tgtcatcaac catgtggtgg agtttgatca tgtaacatga ctccgacaga aaacacgggtt 60
 atcaaatggt tcacctgtac gtccatcgta aaggattgtt ttggcategc tatccataacc 120
 agcttcttta acagttgacc aaagatcttca agaacttgct ccgtcaaaga ctgggtttgc 180
 gatgtgaata ccaagagtac gagctgccat cccaaagggtgg agttccataa cctgaccgat 240
 attcatacgt gatggcaccc caagtgggtt caacatgata tcgactggag ttccatctgg 300
 aaggtaaggc atatcttcta caggaacgat acgagagaca acccctttat ttccgtgacg 360
 tccggccatc ttatctccga ctttgatctt acgttttga gcgtatgtaga cgcgaaaccag 420
 catgttgaca cctgattgca attcatctcc atttgcacgt gtaaagatct taacatcacg 480
 aaccacacca tcagctccgt gtggcacacg aagagaagtg tcacgtactt cacgagattt 540
 atctccgaag atagcgtgca agagccgttc tttagctgaa aggtctttctt caccctttagg 600
 tgttacttta cctacaagga tatccccttc tttaacctca gcaccgatac ggataatacc 660
 catttcgtca agatcttaa gggcatcttc cccaaacgttt gggatttcac gagtaatttc 720

ttcaggtcca . 730

<210> 16
<211> 697
<212> DNA
<213> *Streptococcus equinus*

<400> 16
cactcgcaa attccaaacg ttggtaaga agctctaaa gaccttgcac aaatgggtat 60
tatccgtatc ggtgctgaag ttaaagaagg tgacatcctt gtaggtaaag taacacctaa 120
aggtgaaaaaa gacctttctg ctgaagagcg ctttcttac gcaatctcg gtgataaattc 180
acgtgaagtt cgtgatacat cactcgtgt accacacggt ggagatggtg tcgttcgtga 240
cgtaaaaatc tttacacgtg caaacggta tgaattacaa tcaggtgtta acatgctcgt 300
tcgtgtttat atcgacacaaa aacgtaaaat caaagtccga gataaaatgg ccggcgtca 360
cggttaacaaaa ggggttgttt ctctgttgc tccagttgaa gacatgcctt atcttccaga 420
cggaactcca gtcgatatac ttttgcaccc acttgggtg ccatctcgta tgaacatcgg 480
acaagttatg gagtttcacc ttggatggc tgctcgtaac cttggattc acattgcaac 540
accagtcttt gatggggcaa cttctgaaga cttttggat acagttacg aagctggat 600
ggctagcgcac gctaagacag ttcttacga tggacgtact ggtgaaccat ttgataaccg 660
tgtgtcagtt ggtgtcatgt acatgattaa acttcac 697

<210> 17
<211> 731
<212> DNA
: .
<213> *Streptococcus constellatus*

<400> 17
agttgtcatc aaccatgtgt gcaatttaat catatacatg acacccgacag atacacggtt 60
gtcaaacggc tcgcccgtac gaccatcata aagaatcgac ttggcatcgc tatccatgcc 120
tgcttcacga acagtatccc aaaggtcatc tgagcttgct ccgtcaaata ctggcggtgc 180
tatgtggata ccaaggttgc gagcagecat accaaggtga agctccataa cctgtccgat 240
attcatacgt gatggcaccc caagtgggtt caacatgatg tctactggtg ttccgtctgg 300
aagataaggc atatcctcaa ctggAACGAT acgggaaaca acccctttat ttccgtggcg 360
tccggccatc ttatccccaa cgcggatctt tcgttttga gcaatgtaaa cacgcaccaa 420
catgttgaca ccagattgca attcatcacc gttcgcacga gtaaagattt tcacatcacg 480
gacaacccca gcaccaccat gtggtaacacg aagagatgtg tcacgtactt cacgagattt 540

atcaccgaaa attgcatgaa gcaggcggttc ttca	gcccggat aagtctttt caccttcgg	600
cgttacttta ccgacaagaa tgcgcctc tttcacctca	gcaccaatgc ggataattcc	660
catttcgtca aggtctctta ggcgcatttc cccaa	cggttt ggaatttcgc gcgtatttc	720
ttcaggtcca a		731

<210> 18
 <211> 697
 <212> DNA
 <213> *Streptococcus anginosus*

<400> 18		
cacgcgcgaa attccaaacg tcggtaaga tgcttgaga	gacccgtacg aaacggaaat	60
tatccgcatt ggtgctgagg taaaagaagg cgacattctt	gtcggtaaag taacaccgaa	120
aggtgaaaaaa gacttatctg ctgaagaacg cctgcttcat	gcaattttcg gtgataaattc	180
tcgtgaagta cgtgataactt cccttcgtgt accacatggt	ggtgcagggg ttgtccgtga	240
tgtgaaaatc tttactcgtg cgaacggtga tgaattgcaa	tctgggtca acatgttggt	300
acgtgtttac atcgctcaaa aacggaaaaat ccgtgttggg	gataagatgg ctggacgtca	360
cggaaacaaa ggggttgttt cccgcattgt tccagtttag	gatatgccgt atcttccaga	420
tggaacacca gttgatatta tggtaaccc acttgggtg	ccatctcgta tgaatattgg	480
tcaagttatg gagcttcacc tcggatggc tgctcgcaac	cttggcattc acattgcaac	540
accagtattt gacggggcta gtcagatga tctttggaa	accgttcgtg aagctggcat	600
ggatagcgat gctaagacaa tccttatga tggccgtact	ggtgagccat ttgataatcg	660
tgtatccgtt ggtgtcatgt acatgatcaa actccac		697

<210> 19
 <211> 728
 <212> DNA
 <213> *Streptococcus dysgalactiae*

<400> 19		
tgtcatcaac catgtggtgg agttaatca tgtacatgac	accaacggat acacggttgt	60
caaattggttc gccagtacgt ccatcataaa ggaccgtctt	agcatcgcta tccataaccag	120
cttcacgaac agtgtccaa aggtcttctg atgaagcccc	gtcaaagaca ggtgttgcaa	180
tgtgaatacc aagattacga gcagccatac caaggtgaag	ttccataacc tgaccaatgt	240
tcatccgtga tggcacccca agagggttca acatgatgtc	aactgggttt ccatctggaa	300
ggtatggcat gtcttcaact ggtacaatac gtgaaacgac	accctgttt ccgtgacgac	360

cagccatttt atctccgact ttgatcttac gttttgagc aatgtaaaca cgcacaagca	420
tattaacacc tgattgcaat tcatcgccgt tagcgcgtgt aaagattttc acatcacgaa	480
cgataccatc accaccgtga ggtacacgaa gggacgtatc acgaacttca cgtgatttat	540
ctccaaagat ggcatgcaag agacgctttt cagcagaaag gtcttttca cctttaggtg	600
tgactttacc tactaagatg tcgccttctt taacctcagc accgatacgg ataattccca	660
tttcgtcaag gtctttgagc gcttcttac caacgtttgg aatttcgcgg gtgatttctt	720
caggtcaa	728

<210> 20	
<211> 728	
<212> DNA	
<213> <i>Streptococcus bovis</i>	
<400> 20	
tgtcatcaac catgtggta agtttgcata tgtacatgtt accaacagag acacgattat	60
caaatggttc acctgtacga ccgtcataaaa gaactgtctt agcgtcgcta tccataccag	120
cttcacgaac agtatccaa aggtcttctg aagttgcccc gtcaaagact ggagttgcaa	180
tgtgaataacc gaggttacga gctgccatac caaggtgaag ttccataact tgtccgatat	240
tcatacgaga tggcacccca agagggttca acatgatatac aactggagtt ccgtctggaa	300
gatatggcat gtcttcaaca ggaacgatac gagaaacaac cccttgttt ccgtgacgac	360
cggccatttt atctccgact ttgattttac gttttgtgc aatgtaaaca cgaacgagca	420
tgttgcacacc tgattgcaat tcacccgt tagcacgtgt gaagattttt acatcacgaa	480
caacaccgtc tccaccgtgt ggcacacgaa gtgatgtatc acgtacttca cgagatttat	540
caccgaagat tgcgtgaaga aggcgttctt cagcagaaag gtcttttca cctttaggtg	600
ttactttacc tacaaggata tcacccctt taacctcagc accgatacgg ataataccca	660
tttcgtcaag gtctttaaga gcttcttac caacgtttgg aatttcgcga gtgatttctt	720
caggtcaa	728

<210> 21	
<211> 728	
<212> DNA	
<213> <i>Streptococcus acidominimus</i>	
<400> 21	
ttgtcatcaa ccatgtggtg gagcttaatc atgtacatga caccaacaga cacacggta	60
tcaaatggtt caccagtacg accatcataa agaatcgaaa tagcatcgct gtccattcct	120

gcctctttaa cagttgacca gagatcctct gagctcgac catcgaaaac cggtgttgcg	180
atatggatac ccaagttacg agcagccata cccaagtgcg gttccataac ctgaccaata	240
ttcatacgag atggcaccgc aagtgggttc aacatgtatgt caactgggtgt tccatctgga	300
agatatggca tgtcttcaac tggtacaata cgagaaacga caccctgtt accgtgacga	360
ccggccatct tatctccgac cttaatcttgcgtt cgttttttagt cgatatacac acgtaccagc	420
atattaacac cagactgttag ctcatcacca ttagcacgcg taaagattt cacatcacga	480
acaacaccat ctgcaccgtg tggcacacgt agagaggtat cacgtacttc acgtgatttg	540
tcaccgaaga tagcatgcaa gagacgctcc tcagcagaaa gatcttttc accttttggt	600
gtcaccttac caacaagaat atgccttct ttaacttctgcgtt caccgatacg gataatacc	660
atttcgtcaa ggtcttttagt ggcttcttca ccaacgtttg gaatttcacg agtaatttct	720
tcaggtca	728

<210> 22
 <211> 733
 <212> DNA
 <213> *Streptococcus agalactiae*

<400> 22	
tgagttgtca tcaaccatgt ggtgaagttt gatcatgtac atgacacccaa ctgacacacacg	60
gttatcgaat ggttcaccag tacgaccatc ataaagaaca gtcttagcat ctgaatccat	120
acctgcttct tgaacagttt cccaaaggtc ttctgaagaa gccccatcaa agactggcgt	180
tgcaatatga atacctaaat tacgagcagc catacctaaa tgaagctcca taacttgc	240
gatattcata cgtgatggca ccccaagtgg gttcaacatg atatcaactg gcggtccatc	300
tggtaagtaa ggcatatctt caacaggaac aatacgttag acgacacccctt tgttccgtg	360
acgaccggcc atcttatac acgtttgtat ttacgtttt tgagcgatat aaacgcggac	420
aagcatatta acacctgatt gcaattcatc accatggca cgagtaaaga ttttaacgtc	480
acgaactact ccatcgccac cgtgaggtac acgttagtgaa gtatcacgaa cttcacgtga	540
tttatcacca aaaatggcat gcaagagacg ttcttcagca gataagtcct tttcacccctt	600
aggtgttacc ttaccaacaa gaatgtcacc ttctttacc tcagcaccaaa tgccgataat	660
tcccatatca tcgagatcac gtagtgaatc ttcaccaaca tttggattt cacgagtaat	720
ttcttcaggt cca	733

<211> 714

<212> DNA

<213> *Streptococcus difficile*

<400> 23

ttgtcatcaa ccatgtggtg aagtttgc atgtacatga caccaactga cacacggta 60
tcgaatggtt caccagtatg accatcataa agaacagtct tagcatctga atccatacct 120
gcttcttcaa cagttccca aaggctttct gaagaagccc catcaaagac tggcggtgca 180
atatgaatac ctaaattacg agcagccata cctaaatgaa gctccataac ttgtccgata 240
ttcatacgtg atggcaccctt aagtgggttc aacatgatata caactggcgt tccatctgg 300
aaataaggca tatcttcaac aggaacaata cgtgagacga cacctttgtt tccgtgacga 360
ccggccatct tatcaccgac tttgattttt cgtttttgag cgatataaac gcggacaagc 420
atattaacac ctgattgcaa ttcatacaccat tttgcacgag taaagatttt aacgtcacga 480
actactccat cgccaccgtg aggtacacgt agtgaagtat cacgaacttc acgtgattta 540
tcaccaaaaa tggcatgcaa gagacgttct tcagcagata agtcctttt acccttaggc 600
gttaccttac caacaagaat gtcaccttct tttacctcag caccaatgcg gataattccc 660
atttcatacga gatcacgtag tgaatcttca ccaacatttg gaatttcacg agta 714

<210> 24

<211> 728

<212> DNA

<213> *Streptococcus intermedius*

<400> 24

tgtcatcaac catgtggta agcttaatca tgtacatgac accaacggac acacggttat 60
caaacggttc gccagtagt ccatcataaa ggattgtctt agcatcgcta tccatacctg 120
cttcacgaac ggtttccaa agatcatctg agctagctcc gtcaaagact ggcgttgc 180
tgtggatacc aaggttgcga gcagccatac cgaggtgcaa ttccataact tgtccgatata 240
tcatacgtga cggcaccctt agaggattca acatgatatac aactggtgac ccgtctggaa 300
gatacggcat atcctcaact ggaacaatgc gggaaacaac ccctttgtt ccgtggcg 360
cggccatctt atctccaacg cggattttcc gttttgagc gatataaaca cgtaccaaca 420
tgttgacacc ggattgcaat tcatcaccgt tcgcacgagt aaagattttt acatcacgga 480
caacacctgc accaccgtgt ggtacacgaa gggaggtatc acgcacattca cgagacttat 540
caccaaaaaat tgcatgaagc aggcgttctt cagcggataa atcttttca ccttcggcg 600
ttactttacc gacaagaatg tcgccttctt ttacctcagc accaatgcgg ataattccca 660

23/32

tctcgtaag gtctctcaaa gcatcttccc cgacgtttg aatttcgcgc gtgatttctt 720
 caggtcca 728

<210> 25
 <211> 728
 <212> DNA
 <213> *Streptotoccus equi*

<400> 25
 tgtcatcaac catgtggta agcttaatca tatacatgac accaactgac acacgattat 60
 caaacggctc accagtacgg ccatcataaa gaacagtctt agcatcgcta tccatacctg 120
 cttcacgaac agtttccaa aggtcctcag acgttagctcc gtcaaagacc ggtgttgcga 180
 tatggatacc caaattacga gcagccatac ctaggtgaag ctccataacc tgtccaatgt 240
 tcatacgaga cggcacccca agagggttca gcatgatgtc aacaggggtt ccgtctggca 300
 gatatggcat atcctcaacc ggtacaatac gtgagacgac acccttgtta ccatgacgccc 360
 cggccatttt atctccgacc ttgatttac gctttgagc aatgtaaaca cgcaccagca 420
 tattaacacc tgattgaagc tcatcaccat ttgcgcgtgt aaagatcttc acatcacgta 480
 caatcccgtc accaccatga ggaacacgta acgaggtatc acgaacctca cgtgatttat 540
 caccaaagat agcatgcagg agacgttctt cagcagaaag gtcttttca cccttaggag 600
 ttaccttacc aacaagaata tcgccttcct tgacctctgc accgatacgg ataataccca 660
 tttcatcaag gtccttgagg gcttcttcac caacgtttg cacttcacgt gtgatttctt 720
 caggtcca 728

<210> 26
 <211> 697
 <212> DNA
 <213> *Enterococcus gallinarum*

<400> 26
 cactcgtaa atcccgaaatg tcgggaaaga cgcattgaaa gatctagacg aaatgggtat 60
 catccgcatt ggtgcggaag tcaaagatgg cgatctgtt gttggtaaag taacgcctaa 120
 aggggtaacg gaactatctg cagaagaacg cttgcttcat gcaatcttg gtaaaaaagc 180
 ccgcgaagtc cgcgatactt ctctgcgcgt acctcacggt ggtggcggaa tcgtccatga 240
 tgtgaaaatc tttacccgcg aagctggcga tgaattgtca ccaggtgtca atatgctgt 300
 tcgcgtgtat atcgaaat aacggaaaat ccatgaaggg gataaaatgg ccggccgtca 360
 cgaaaaataaa ggggtcgaaaat ctcgcattat gccagaagaa gacatgcctt tcttaccaga 420

24/32

cggtagcacca	gttgatata	ttttgaaccc	attaggggtg	ccttcacgga	tgaacattgg	480
acaagtattg	gaattacact	taggaatggc	tgcccgccaa	ttaggaatcc	acgtggctac	540
accagtcttt	gatggtgcca	gcgatgaaga	tgtctggca	acagttgcag	aagccggcat	600
ggctagcgac	gccaaaaccg	ttttgtatga	tggccgtact	ggagaaccat	ttgatggtcg	660
aatctccgta	ggtgtcatgt	atatgatcaa	attggcc			697

<210> 27

<211> 727

<212> DNA

<213> *Enterococcus casseliflavus*

<400> 27

tgtcatcaac	catgtggcc	aatttgcata	tgtacatgac	accaacggag	atgcggccat	60
caaatggttc	gccggtacgt	ccgtcgtaaa	gcactgtttt	ggcatcgctg	gccattcctg	120
cttcagcaac	cgttgcacaa	acatcttcat	cgctggctcc	atcaaagact	ggtgttgcca	180
cgtgaatgcc	taattgacgc	gcagccattc	ctaagtgtaa	ctctaatact	tgtccaatgt	240
tcatccgaga	aggtaccct	aatgggttca	gcatgatatac	gactgggttg	ccatctggta	300
agaaaggcat	gtcttcttct	ggcataatgc	gagaaacgac	ccctttgttt	ccgtgacgtc	360
cggccatttt	atccccttca	tggattttcc	gttttgaac	gatataaacg	cgaaccagca	420
tgttcacacc	tggtgacaat	tcatcgccag	cttcgggtt	aaagattttg	acatcggtt	480
cgattccgccc	gccggcggtga	ggcacgcgt	gagaagtgtc	acgcacttcg	cgggcttttt	540
caccaaagat	tgcgtgcaac	aaacgctttt	ctgctgaaag	ttccgttacc	cctttggcg	600
tgactttccc	aacaaggaga	tcgccccat	tgacttccgc	accaatgcgg	ataatgccc	660
tttcgtctag	gtcttcaac	gcgtcttccc	aacgttcggg	atttcgcgag	tgatttcttc	720
aggtcca						727

<210> 28

<211> 721

<212> DNA

<213> *Enterococcus saccharolyticus*

<400> 28

tgtcatcaac	catgtggca	agtttaatca	tgtacattac	cccaacagag	atacgaccat	60
cgaatggttc	accgtacgt	ccgtcataaa	gaacagttt	cgcacgcgc	gccatgccc	120
cttcgcgaac	tgtttccat	acgtcatcat	ctgatgcacc	atcaaatact	ggtgttagcta	180
catggatgcc	taactgacgt	gcagccatcc	ctaagtgtaa	ttccaatact	tgtccgatgt	240

25/32

tcatacgaga tggtaactcct agtgggttca acatgatatac aactggtgta ccgtctggta	300
agaatggcat gtcttcttct ggcataatgc gagagacaac cccttgcata ccatgacgtc	360
ccgcccattt atctccttcg tgaatcttac gttttgcac gatataaaca cgaactaaca	420
tgttcacacc tggagataat tcgtcgctg cttcacgggt aaagattttt acatcgtaaa	480
cgataaccgcc accggccgtga ggaacacgta atgatgtatc acgtacttca cgtgctttt	540
caccgaagat tgcgtgcaat agacgttctt ctgcagataa ttcggttacc cctttaggag	600
tgactttacc tactaataag tcgccatctt gtacttcggc accgatacgg ataataacca	660
tttcgtctaa gtcttttaat gcgtcttccc caacgttagg aatttcgctg gtattttca	720
g	721

<210> 29
 <211> 727
 <212> DNA
 <213> Enterococcus faecium

<400> 29	
tgtcatcaac catgtgagca agtttgatca tgtacatcac accgacagac acacgtccat	60
caaattggttc acctgtacgt ccgtcgtaa gaacagtttt cgcatcgctg gccataccgg	120
cttcacgaac tgtttccat acgtcttcat cacttgaccat atcaaataact ggctttgcta	180
cgtggataacc taactgacgt gcagccatac ccaagtgtaa ttccaaataact tgcccgatgt	240
tcatacgta aggcacccct aaaggattca gcatgatatac gattgggttt ccatcaggtt	300
ggaatggcat atctttttcc ggcataatac gggatacaac ccctttatcc ccgtgacgac	360
cggccatttt atccccttca tggattttac gttttgaac gatataaaca cgaactaaca	420
tgtttacgcc tggtgacaat tcattccag cttcacgagt aaagattttc acatcgtaaa	480
cgataaccgcc gcccgcgtgt ggtacacgta atgatgtatc gcggacttca cgagctttt	540
cgc当地ggat cgcataatgc agacgttctt ctgcagataa ttctgttacc cctttggcg	600
tgactttccc tacaagcaaa tcgccatctt ggacttctgc accaatacgg atgataacca	660
tttcgtctaa atcttttaat gcgtcttccc gacatttaggg aatttcgctg tgatttttc	720
aggcata	727

<210> 30
 <211> 725
 <212> DNA
 <213> Enterococcus faecalis

<400> 30

26/32

tgtcatcaac catgtggct aattnaatca tatacatgac accaacggaa atacggttat 60
 caaatggttc acctgtacgt ccatcgtaaa gaactgtttt agcatcgcta gccataccag 120
 ctacacgaac agttcccaa acgtcttcat cggttgcccc atcgaaaaca ggtgttgcga 180
 cgtgaataacc taattggcga gcagccatac ctaagtgtaa ttcaagtact tgtccgatat 240
 tcatacgaga aggtacccct aatgggttca acatgatatc aacaggtgtt ccgtcaggta 300
 agaatggcat atcttcttcc ggcataatac gggaaacaac ccctttattt ccgtgacgtc 360
 ccgcatttt atctccttcg tgaattttac gttttgaac gatatagaca cgaactaaca 420
 tggtgacacc tggtgataat tcatcgccag ctacgagt aaagatttc acatcatgaa 480
 cgataccgccc gccaccgtga ggtacacgga gagacgtatc acgaacttcg cgggctttt 540
 cccccgaagat tgcgtgtaat aaacgttctt ctgcagataa ttctgtgacc cctttaggtg 600
 tgactttccc aactagtaag tcgccatctt gaacttcagc accaatgcgg ataatcccc 660
 tttcgctcaa gtcttcaac gcgtcttccc aacgtttgga atttcacggg tatttcttca 720
 ggtca 725

<210> 31
<211> 570
<212> DNA
<213> *Enterococcus avium*

<400> 31
gtccatcata aagaacggtc ttagcatctg ctgcatacg agttcacga actgtttccc 60
aaacatcgct atcttgcga ccatcgaa ctgggtgtcgc aacatggata cctagttggc 120
gagccgccc tcccaagtgt aattccaaaca cttgtccgat gttcatccga gatggcacac 180
ctaattgggtt caacatgata tcaactggcg taccgtctgg taagaaaggc atgtttctt 240
ctggcataat gcgagaaaacg acccccttat ttccgtgacg gccggccatt ttatcccctt 300
catgaatctt acgttttgc acgatgtaca cgccactaa catattaca cctggagata 360
attcatcgcc tgcttacga gtaaagatct tcacatcgat aacgatcccg ccggcaccat 420
gcggtaacacg aagagatgta tcacgaactt cacgagcctt ttcaccaaag atcgcatgca 480
acaaacgttc ttcaagctgat aattctgtta cccctttagg agtgacttta ccaactaata 540
aatcaccatc atgaacttca gcaccaatac 570

<210> 32
<211> 732
<212> DNA
<213> *Abiotrophia defectiva*

<400> 32

gaagttgtca tcaaccatgt gggccaactt aatcatgtac ataaccccaa cagagacttt 60
 acggtaaat ggttcaccgg ttcgaccatc atataagata gtcttagcgt cagttctaa 120
 gccggcttcc ttaactgttt cccagacatc ttcttcacta gcaccgtcaa agacaggtgt 180
 tgcaatcttg atgcccattt cgcgagcagc catccccaaag tgtaactcta ggacttgccc 240
 gatgttcata cgggatggaa cccctaattgg gttcaacatg atatcaactg gggtaccatc 300
 tggtaagaat ggcataatctt cttccggcat gataagggag acaacccctt tgttaccgtg 360
 acgaccggcc atcttatccc cttcattgtat tttacgtttt tgtacgtgt agacgcccac 420
 tagcttggtg acacctggtg ccaattcgtc gccagcttcg cggtaaaga ttttaacgtc 480
 gtggacaatc ccgcggccgc cgtgtggcac acgcaaggaa gtatcacgta cttcacgcgc 540
 cttctcaccg aagatagcat ggagcaagcg ttcttccgca gacaactcgg tcacaccttt 600
 tgggttacc ttaccaacta agatatcgcc gtctttact tccgcggca tacagataat 660
 cccgtcttgg tctaagtact tgagggcatc ttcggacacg tttgaaattt cgcgtgtaat 720
 ttcttcaggt ca 732

<210> 33

<211> 727

<212> DNA

<213> *Gemella morbillorum*

<400> 33

tgtcatcaac catgtgtgca agtttatcat gtacattacc cctacagata cacggctatc 60
 aaatggctca cctgtacgtc cgtcataaag aactgtctta gcatcttttag ccattccagc 120
 ttccgcaact gtagaccaaa catcttcatc agtagcacca tcgaataactg gtgtagctac 180
 gtggattcca agttgttttag cagccatacc taagtgtac tctaataactt gtccaatgtt 240
 catacgagat ggaaccccaa gtgggtttaa cattacgtca actgggttac catctggtag 300
 gtaaggcata tcttcttctg gtaagatatt tgagataacc ctttggtag cgtgacgacc 360
 ggccatttta tctcctacac gaattttacg tttttggacg ataaatacac gaacaagttc 420
 atttacaccg ttaggttaatt cagcaccatc ttcacgttta aagattttaa catcagcaac 480
 tactccatca gcaccgttag gtacacgtaa tgaagtatca cgtacttctt tagatttagc 540
 tccaaagata gcatataata atttttcttc tggagttgt tcagttatc ctttcgggtgt 600
 aactttacct actaaaaat atccatctt aacttcagcc ccaatacgaa tgattcctcg 660
 tgcatctaag tttctaagtg cattttcacc ctacgttgg aatctcacga gtaatttctt 720

caggtca	727
<210> 34	
<211> 726	
<212> DNA	
<213> <i>Gemella haemolysans</i>	
<400> 34	
tgtcatcaac catgtgtgca agtttaatca tgtacattac ccctacagat acacggctat	60
caaatggctc acctgtacgt ccgtcataaaa gaactgtctt agcatcttta gccattccag	120
cttccgcaac tgttagaccaa acatcttcat cagtagcacc atcgaatact ggttagcta	180
cgtggattcc aagttgttta gcagccatac ctaagtgtag ctctaatact tgtccaatgt	240
tcatacgaga tggaaacccc agtgggttta acattacgtc aactgggtta ccatctggta	300
ggtaaggcat atcttcttct ggtaagatat ttgagataac ccctttgttta ccgtgacgac	360
cggccatttt atctcctaca cgaattttac gttttggac gataaataca cgaacaagtt	420
catttacacc gtttaggtaat tcagcaccat ctacgtttt aaagattttta acatcagcaa	480
ctactccatc agcaccgtga ggtacacgta atgaagtatc acgtacttct ttagatttag	540
ctccaaagat agcatataat aatttttctt ctggagtttg ttcagttaat ctttcgggt	600
taactttacc tactaaaata tctccatctt taacttcagc cccaaatacga atgattcctc	660
gtgcacatctaa gtttotaagt gatatttac ctagtttgg aatctcacga gtattcttca	720
ggtcca	726
<210> 35	
<211> 719	
<212> DNA	
<213> <i>Granulicatella adjacens</i>	
<400> 35	
catcaaccat gtgagcaagt ttgatcatgt acataacccc tactgacaca cggttatcga	60
atggttcccc tgtacgtcca tcatatagaa ttgtttcgc atcacgagcc atacccgctt	120
ctgcaacagt tccccatacg tcttcatctt gcgcaccatc gaatactggt gttgcgtatgt	180
aaataacctaa ttacgagca gccatcccta agtgtaactc taacacttgt ccgatgttca	240
tacgtgaagg taccctaat gggtaaca ttagtcaac tgggtttcca tctggtaaga	300
atggcatatc ttcttccggc ataatacggg aaacaacccc tttattaccg tgacgtccgg	360
ccatcttatac cccttcattt atttacgtt tttgtacaat atatacacga actaatttgt	420
ttacgcccagg tgcttaat tcacctgctg cacgtgtgaa tacacgtaca tcacggacaa	480

taccgccacc gccgtgaggt acacgttagag atgtgtcacg aacttcacga gcttttcac 540
cgaagattgc gtgtataaaa cgttcctctg gtgattgttc tgttaaccct ttaggagtt 600
cttaccaac taagatgtca ccatcttaa cttcggcacc gatacgaata attccgtctg 660
cgtcttaggtt cttcaatgcg tcttcccaac gtttggaaatc tcacgagtaa ttcttcagg 719

<210> 36
<211> 21
<212> DNA
<213> amorce

<400> 36
agacggacct tctatggaaa a

21

<210> 37
<211> 20
<212> DNA
<213> amorce

<400> 37
ggacacatac gaccatagtg

20

<210> 38
<211> 21
<212> DNA
<213> amorce

<400> 38
gttgtaacct tcccaawgtca t

21

<210> 39
<211> 20
<212> DNA
<213> amorce

<400> 39
gtcttcwtgg gygatttccc

20

<210> 40
<211> 21
<212> DNA
<213> amorce

<220>
<221> misc_feature
<222> (8)..(8)
<223> n représente i

<400> 40

30/32

accgtggngc wtggtrgaa t

21

<210> 41
<211> 20
<212> DNA
<213> amorce

<400> 41
aaccaattcc gyatyggtt

20

<210> 42
<211> 20
<212> DNA
<213> amorce

<220>
<221> misc_feature
<222> (3)..(3)
<223> n représente i

<400> 42
agngggttta acatgatgtc

20

<210> 43
<211> 20
<212> DNA
<213> amorce

<220>
<221> misc_feature
<222> (3)..(3)
<223> n représente i

<400> 43
agngcccaaa cctccatctc

20

<210> 44
<211> 21
<212> DNA
<213> amorce

<400> 44
ctccaaatgtga acagatgtgt a

21

<210> 45
<211> 26
<212> DNA
<213> amorce

<400> 45
ttacccaaact taattgagat tcaaac

26

<210> 46
<211> 22
<212> DNA
<213> amorce

<400> 46
agtatttatg ggtgatttcc ca

22

<210> 47
<211> 26
<212> DNA
<213> amorce

<400> 47
ggacgttata aaatcaacaa aaaatt

26

<210> 48
<211> 23
<212> DNA
<213> amorce

<400> 48
agttataacc atcccaagtc atg

23

<210> 49
<211> 23
<212> DNA
<213> amorce

<400> 49
tgaagtttat catcaaccat gtg

23

<210> 50
<211> 18
<212> DNA
<213> amorce

<400> 50
cccaaaaacgt tgtccacc

18

<210> 51
<211> 20
<212> DNA
<213> amorce

<400> 51
aaccaagcyc ggtaggrat

20

<210> 52
<211> 25

<212> DNA
<213> amorce

<220>
<221> misc_feature
<222> (15)..(15)
<223> n représente i

<400> 52
atgttgaacc cactnggggt gccat

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.